

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number
WO 2004/046340 A2

(51) International Patent Classification⁷:

C12N

(21) International Application Number:

PCT/US2003/037047

(22) International Filing Date:

14 November 2003 (14.11.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/426,095 14 November 2002 (14.11.2002) US

(71) Applicant (for all designated States except US): GENEN-TECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): KRUMMEN, Lynne [US/US]; 3030 22nd Avenue, San Francisco, CA 94132 (US). SHEN, Amy [US/US]; 1828 Parrott Drive, San Mateo, CA 94402 (US).

(74) Agents: STEFFES, David et al.; Sidley Austin Brown & Wood LLP, 1501 K Street, N.W., Washington, D.C. 20005 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TI, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 2004/046340 A2

(54) Title: INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-EXPRESSING PRODUCTION CELL LINES

(57) Abstract: This invention relates to a DNA construct, methods of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest. In one method, stable clones capable of producing a high level of a product of interest are generated from one step of a direct selection immediately after transfection.

BEST AVAILABLE COPY

INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
EXPRESSING PRODUCTION CELL LINES

This application claims priority under 35 U.S.C. § 119(e) from U.S. provisional application serial no. 60/426,095, filed November 14, 2002, which is herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to a DNA construct, a method of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest.

Description of Background and Related Art

The discovery of methods for introducing DNA into living host cells in a functional form has provided the key to understanding many fundamental biological processes, and has made possible the production of important proteins and other molecules in commercially useful quantities.

Despite the general success of such gene transfer methods, several common problems exist that may limit the efficiency with which a gene encoding a desired protein can be introduced into and expressed in a host cell. One problem is knowing when the gene has been successfully transferred into recipient cells. A second problem is distinguishing between those cells that contain the gene and those that have survived the transfer procedures but do not contain the gene. A third problem is identifying and isolating those cells that contain the gene and that are expressing high levels of the protein encoded by the gene.

In general, the known methods for introducing genes into eukaryotic cells tend to be highly inefficient. Of the cells in a given culture, only a small proportion take up and express exogenously added DNA, and an even smaller proportion stably maintain that DNA.

Identification of those cells that have incorporated a product gene encoding a desired protein typically is achieved by introducing into the same cells another gene, commonly referred to

as a selectable gene, that encodes a selectable marker. A selectable marker is a protein that is necessary for the growth or survival of a host cell under the particular culture conditions chosen, such as an enzyme that confers resistance to an antibiotic or other drug, or an enzyme that compensates for a metabolic or catabolic defect in the host cell. For example, selectable genes commonly used with eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (hyg), dihydrofolate reductase (DHFR), thymidine kinase (tk), neomycin resistance, puromycin resistance, glutamine synthetase, and asparagine synthetase.

The method of identifying a host cell that has incorporated one gene on the basis of expression by the host cell of a second incorporated gene encoding a selectable marker is referred to as cotransfection (or cotransfection). In that method, a gene encoding a desired polypeptide and a selection gene typically are introduced into the host cell simultaneously. In this case of simultaneous cotransfection, the gene encoding the desired polypeptide and the selectable gene may be present on a single DNA molecule or on separate DNA molecules prior to being introduced into the host cells. Wigler *et al.*, Cell, 16:777 (1979). Cells that have incorporated the gene encoding the desired polypeptide then are identified or isolated by culturing the cells under conditions that preferentially allow for the growth or survival of those cells that synthesize the selectable marker encoded by the selectable gene.

The level of expression of a gene introduced into a eukaryotic host cell depends on multiple factors, including gene copy number, efficiency of transcription, messenger RNA (mRNA) processing, stability, and translation efficiency. Accordingly, high level expression of a desired polypeptide typically will involve optimizing one or more of those factors.

For example, the level of protein production may be increased by covalently joining the coding sequence of the gene to a "strong" promoter or enhancer that will give high levels of transcription. Promoters and enhancers are nucleotide sequences that interact specifically with proteins in a host cell that are involved in transcription. Kriegler, Meth. Enzymol., 185:512 (1990); Maniatis *et al.*, Science, 236:1237 (1987). Promoters are located upstream of the coding sequence of a gene and facilitate transcription of the gene by RNA polymerase. Among the eukaryotic promoters that have been identified as strong promoters for high-level expression are the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, Rous sarcoma virus long terminal repeat, and human cytomegalovirus immediate early promoter (CMV).

Enhancers stimulate transcription from a linked promoter. Unlike promoters, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter, although in practice enhancers may overlap physically and functionally with promoters. For example, all of the strong promoters listed above also contain strong enhancers. Bendig, Genetic Engineering, 7:91 (Academic Press, 1988).

The level of protein production also may be increased by increasing the gene copy number in the host cell. One method for obtaining high gene copy number is to directly introduce into the host cell multiple copies of the gene, for example, by using a large molar excess of the product gene relative to the selectable gene during cotransfection. Kaufman, Meth. Enzymol., 185:537 (1990). With this method, however, only a small proportion of the cotransfected cells will contain the product gene at high copy number. Furthermore, because no generally applicable, convenient method exists for distinguishing such cells from the majority of cells that contain fewer copies of the product gene, laborious and time-consuming screening methods typically are required to identify the desired high-copy number transfecants.

Another method for obtaining high gene copy number involves cloning the gene in a vector that is capable of replicating autonomously in the host cell. Examples of such vectors include mammalian expression vectors derived from Epstein-Barr virus or bovine papilloma virus, and yeast 2-micron plasmid vectors. Stephens & Hentschel, Biochem. J., 248:1 (1987); Yates *et al.*, Nature, 313:812 (1985); Beggs, Genetic Engineering, 2:175 (Academic Press, 1981).

Yet another method for obtaining high gene copy number involves gene amplification in the host cell. Gene amplification occurs naturally in eukaryotic cells at a relatively low frequency. Schimke, J. Biol. Chem., 263:5989 (1988). However, gene amplification also may be induced, or at least selected for, by exposing host cells to appropriate selective pressure. For example, in many cases it is possible to introduce a product gene together with an amplifiable gene into a host cell and subsequently select for amplification of the marker gene by exposing the cotransfected cells to sequentially increasing concentrations of a selective agent. Typically the product gene will be coamplified with the marker gene under such conditions.

The most widely used amplifiable gene for that purpose is a DHFR gene, which encodes a dihydrofolate reductase enzyme. The selection conditions used in conjunction with a DHFR gene are the absence of glycine, hypoxanthine and thymidine (GHT) with or without the presence of methotrexate (Mtx). A host cell is cotransfected with a product gene encoding a desired protein

and a DHFR gene, and transfectants are identified by first culturing the cells in GHT -free culture medium that may contains Mtx. A suitable host cell when a wild-type DHFR gene is used is the Chinese Hamster Ovary (CHO) cell line deficient in DHFR activity, prepared and propagated as described by Urlaub & Chasin, Proc. Nat. Acad. Sci. USA, 77:4216 (1980). The transfected cells then are exposed to successively higher amounts of Mtx. This leads to the synthesis of multiple copies of the DHFR gene, and concomitantly, multiple copies of the product gene. Schimke, J. Biol. Chem., 263:5989 (1988); Axel *et al.*, U.S. Patent No. 4,399,216; Axel *et al.*, U.S. Patent No. 4,634,665. Other references directed to co-transfection of a gene together with a genetic marker that allows for selection and subsequent amplification include Kaufman in Genetic Engineering, ed. J. Setlow (Plenum Press, New York), Vol. 9 (1987); Kaufman and Sharp, J. Mol. Biol., 159:601 (1982); Ringold *et al.*, J. Mol. Appl. Genet., 1:165-175 (1981); Kaufman *et al.*, Mol. Cell Biol., 5:1750-1759 (1985); Kaetzel and Nilson, J. Biol. Chem., 263:6244-6251 (1988); Hung *et al.*, Proc. Natl. Acad. Sci. USA, 83:261-264 (1986); Kaufman *et al.*, EMBO J., 6:87-93 (1987); Johnston and Kucey, Science, 242:1551-1554 (1988); Urlaub *et al.*, Cell, 33:405-412 (1983).

To extend the DHFR amplification method to other cell types, a mutant DHFR gene that encodes a protein with reduced sensitivity to methotrexate may be used in conjunction with host cells that contain normal numbers of an endogenous wild-type DHFR gene. Simonsen and Levinson, Proc. Natl. Acad. Sci. USA, 80:2495 (1983); Wigler *et al.*, Proc. Natl. Acad. Sci. USA, 77:3567-3570 (1980); Haber and Schimke, Somatic Cell Genetics, 8:499-508 (1982).

Alternatively, host cells may be co-transfected with the product gene, a DHFR gene, and a dominant selectable gene, such as a *neo*^r gene. Kim and Wold, Cell, 42:129 (1985); Capon *et al.*, U.S. Pat. No. 4,965,199. Transfectants are identified by first culturing the cells in culture medium containing neomycin (or the related drug G418), and the transfectants so identified then are selected for amplification of the DHFR gene and the product gene by exposure to successively increasing amounts of Mtx.

As will be appreciated from this discussion, the selection of recombinant host cells that express high levels of a desired protein generally is a multi-step process. In the first step, initial transfectants are selected that have incorporated the product gene and the selectable gene. In subsequent steps, the initial transfectants are subject to further selection for high-level expression of the selectable gene and then random screening for high-level expression of the product gene. To identify cells expressing high levels of the desired protein, typically one must screen large numbers

of transfectants. The majority of transfectants produce less than maximal levels of the desired protein. Further, Mtx resistance in DHFR transformants is at least partially conferred by varying degrees of gene amplification. Schimke, Cell, 37:705-713 (1984). The inadequacies of co-expression of the non-selected gene have been reported by Wold *et al.*, Proc. Natl. Acad. Sci. USA, 76:5684-5688 (1979). Instability of the amplified DNA is reported by Kaufman and Schimke, Mol. Cell Biol., 1:1069-1076 (1981); Haber and Schimke, Cell, 26:355-362 (1981); and Fedespiel *et al.*, J. Biol. Chem., 259:9127-9140 (1984).

Several methods have been described for directly selecting such recombinant host cells in a single step. One strategy involves co-transfected host cells with a product gene and a DHFR gene, and selecting those cells that express high levels of DHFR by directly culturing in medium containing a high concentration of Mtx. Many of the cells selected in that manner also express the co-transfected product gene at high levels Page and Sydenham, Bio/Technology, 9:64 (1991). This method for single-step selection suffers from certain drawbacks that limit its usefulness. High-expressing cells obtained by direct culturing in medium containing a high level of a selection agent may have poor growth and stability characteristics, thus limiting their usefulness for long-term production processes Page and Snyderman, Bio/Technology, 9:64 (1991). Single-step selection for high-level resistance to Mtx may produce cells with an altered, Mtx-resistant DHFR enzyme, or cells that have altered Mtx transport properties, rather than cells containing amplified genes. Haber *et al.*, J. Biol. Chem., 256:9501 (1981); Assaraf and Schimke, Proc. Natl. Acad. Sci. USA, 84:7154 (1987).

Another method involves the use of polycistronic mRNA expression vectors containing a product gene at the 5' end of the transcribed region and a selectable gene at the 3' end. Because translation of the selectable gene at the 3' end of the polycistronic mRNA is inefficient, such vectors exhibit preferential translation of the product gene and require high levels of polycistronic mRNA to survive selection. Kaufman, Meth. Enzymol., 185:487 (1990); Kaufman, Meth. Enzymol., 185:537 (1990); Kaufman *et al.*, EMBO J., 6:187 (1987). Accordingly, cells expressing high levels of the desired protein product may be obtained in a single step by culturing the initial transfectants in medium containing a selection agent appropriate for use with the particular selectable gene. However, the utility of these vectors is variable because of the unpredictable influence of the upstream product reading frame on selectable marker translation and because the upstream reading frame sometimes becomes deleted during methotrexate amplification (Kaufman

et al., J. Mol. Biol., 159:601-621 (1982); Levinson, Methods in Enzymology, San Diego: Academic Press, Inc. (1990)). Later vectors incorporated an internal translation initiation site derived from members of the picornavirus family which is positioned between the product gene and the selectable gene (Pelletier *et al.*, Nature, 334:320 (1988); Jang *et al.*, J. Virol., 63:1651 (1989)).

A third method for single-step selection involves use of a DNA construct with a selectable gene containing an intron within which is located a gene encoding the protein of interest. See U.S. Patent No. 5,043,270 and Abrams *et al.*, J. Biol. Chem., 264(24): 14016-14021 (1989). In yet another single-step selection method, host cells are co-transfected with an intron-modified selectable gene and a gene encoding the protein of interest. See WO 92/17566, published October 15, 1992. The intron-modified gene is prepared by inserting into the transcribed region of a selectable gene an intron of such length that the intron is correctly spliced from the corresponding mRNA precursor at low efficiency, so that the amount of selectable marker produced from the intron-modified selectable gene is substantially less than that produced from the starting selectable gene. These vectors help to insure the integrity of the integrated DNA construct, but transcriptional linkage is not achieved as selectable gene and the protein gene are driven by separate promoters.

Other mammalian expression vectors that have single transcription units have been described. Retroviral vectors have been constructed (Cepko *et al.*, Cell, 37:1053-1062 (1984)) in which a cDNA is inserted between the endogenous Moloney murine leukemia virus (M-MuLV) splice donor and splice acceptor sites which are followed by a neomycin resistance gene. This vector has been used to express a variety of gene products following retroviral infection of several cell types.

A method for selecting recombinant host cells expressing high levels of a desired protein was previously described by the applicants in Lucas *et al.*, Nucleic Acid Research, 24, No. 9: 1774-1779 and U.S. Patent No. 5,561,053. That method utilizes eukaryotic host cells harboring a DNA construct comprising a selectable gene (preferably an amplifiable gene) and a product gene provided 3' to the selectable gene. The selectable gene is positioned within an intron defined by a splice donor site and a splice acceptor site and the selectable gene and product gene are under the transcriptional control of a single transcriptional regulatory region. The splice donor site is generally an efficient splice donor site and thereby regulates expression of the product gene using the transcriptional regulatory region. The transfected cells are cultured so as to express the gene encoding the product in a selective medium which may contain an amplifying agent for sufficient

time to allow cells having multiple copies of the product gene, or cells with a single (or multiple) copy of the gene in a chromosomal loci with high transcriptional activity to be identified.

Other fusion expression constructs have been developed. For example, a fusion of green fluorescent protein with the Zeocin-resistance marker construct has been created. Bennet, R.P. *et al.*, *Biotechniques*, 24(3):478-82, 1998 March. Such constructs were used to allow visual screening and drug selection of transfected eukaryotic cells.

In another example, human prothrombin was overexpressed in transformed eukaryotic cells using a dominant bifunctional selection and amplification marker. Herlitschka, Sabine E. *et al.*, *Protein Expression and Purification*, 8, 358-364, 1996 July. In this reference the marker consisted of the murine wild-type dihydrofolate reductase cDNA and the *E. coli* hygromycin phosphotransferase gene fused in frame. The gene of interest is connected, upstream, by the EMCV untranslated region to the fusion marker gene, forming a dicistronic transcription unit.

With the state of the art in mind, it is one object of the present invention to increase the level of homogeneity with regard to expression levels of stable clones transfected with a product gene of interest, by expressing fused selectable markers (i.e. DHFR and puromycin) and a protein of interest from a single promoter.

It is another object to provide a method for selecting stable, recombinant host cells that express high levels of a desired protein product, which method is rapid and convenient to perform, and reduces the numbers of transfected cells which need to be screened. Furthermore, it is an object to allow high levels of single and multiple unit polypeptides to be rapidly generated from clones or pools of stable host cell transfectants.

It is an additional object to provide expression vectors which bias for active integration events (i.e. have an increased tendency to generate transformants wherein the DNA construct is inserted into a region of the genome of the host cell which results in high level expression of the product gene) and can accommodate a variety of product genes without the need for modification.

SUMMARY OF THE INVENTION

Accordingly, the present invention is directed to a DNA construct (DNA molecule) comprising a 5' transcriptional initiation site and a 3' transcriptional termination site, two selectable genes that have been fused into one open reading frame (preferably amplifiable genes) and a product gene provided 3' to the fused selectable genes, a transcriptional regulatory region

regulating transcription of both the fused selectable genes and the product gene, the fused selectable genes positioned within an intron defined by a splice donor site and a splice acceptor site. The splice donor site preferably comprises an effective splice donor sequence as herein defined and thereby regulates expression of the product gene using the transcriptional regulatory region.

In another embodiment, the invention provides a method for producing a product of interest comprising culturing a eukaryotic cell which has been transfected with the DNA construct described above, so as to express the product gene and recovering the product.

In a further embodiment, the invention provides a method for producing eukaryotic cells having multiple copies of the product gene comprising transfecting eukaryotic cells with the DNA construct described above (where the selectable fused genes are amplifiable genes), growing the cells in a selective medium comprising an amplifying agent(s) for a sufficient time for amplification to occur, and selecting cells having multiple copies of the product gene. After transfection of the host cells, most of the transfectants fail to exhibit the selectable phenotype characteristic of the protein encoded by either of the selectable genes, but surprisingly a small proportion of the transfectants do exhibit one or both of the selectable phenotype, and among those transfectants, the majority are found to express high levels of the desired product encoded by the product gene. Thus, the invention provides an improved method for the selection of recombinant host cells expressing high levels of a desired product, which method is useful with a wide variety of eukaryotic host cells and avoids the problems inherent in, and improves upon, existing cell selection technology.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates schematically the construction of the pSV.IPD. The gene for the protein of interest would be inserted at the polylinker site.

Figures 2-1 to 2-4 depict the nucleotide sequence of the pSV.IPUR plasmid used in constructing pSV.IPD (SEQ ID NO 1).

Figures 3-1 to 3-4 depict the nucleotide sequence of the pSV.ID plasmid used in constructing pSV.IPD (SEQ ID NO 2).

Figures 4-1 to 4-4 depict the nucleotide sequence of the pSV.IPD (SEQ ID NO 3).

Figure 5 illustrates schematically the plasmid, pSV.ID.VEGF, used as a control in Example 1.

Figure 6 illustrates schematically the plasmid, pSV.IPD.2C4, used in Example 1 (SEQ ID NO 4).

Figures 7-1 to 7-8 depict the nucleotide sequence of the pSV.IPD.2C4 plasmid used in Example 1.

Figure 8 depicts a FACS analysis of transiently transfected CHO cells with a GTP plasmid in 250ml spinner transfection. FACS analysis was performed 24 hours after transfection.

Figure 9 depicts the expression level of clones from traditional 10nM MTX selection. Cells were transfected with commercial transfection reagent and directly selected in 10 nM MTX. Individual clones were grown in a 96-well plate. Product accumulated for 6 days prior to ELISA.

Figures 10-1 and 10-2 depict the expression level of clones from 25 and 50 nM MTX direct selections, respectively, of SV40-based constructs derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 11 depicts the expression level of clones from 25 nM MTX direct selection of CMV-based construct derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 12 depicts the titer evaluation in Miniferm. Samples were collected every day and submitted to an HPLC protein A assay for titer.

Figure 13-1 to 13-7 depict the nucleotide sequence of the pCMV.IPD.Heterologous polypeptide (HP) plasmid used in Example 3.

Figure 14-1 to 14-8 depicts the nucleotide sequence of the pSV40.IPD.HP plasmid used in Example 3.

Figure 15 illustrates schematically the plasmid, pCMV.IPD.HP, used in Example 3.

Figure 16 illustrates a time line and titer comparison between a traditional selection and direct selection method described in Example 3. Equivalent titers are indicated horizontally across the illustration. For example, the titers for a 200/300nM SV40-plasmid traditional selection, 100nM SV40-plasmid direct selection and 25nm CMV-plasmid direct selection are roughly equivalent.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Definitions:

The "DNA construct" disclosed herein comprises a non-naturally occurring DNA molecule or chemical analog which can either be provided as an isolate or integrated in another DNA molecule *e.g.* in an expression vector or the chromosome of an eukaryotic host cell.

The term "selectable gene" as used herein refers to a DNA that encodes a selectable marker necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. Accordingly, a host cell that is transformed with a selectable gene will be capable of growth or survival under certain cell culture conditions wherein a non-transfected host cell is not capable of growth or survival. Typically, a selectable gene will confer resistance to a drug or compensate for a metabolic or catabolic defect in the host cell. Examples of selectable genes are provided in the following table. See also Kaufman, Methods in Enzymology, 185: 537-566 (1990), for a review of these.

"Fused selectable genes" as used herein refers to a DNA that encodes at least two selectable markers in the same open reading frame and inserted into an intron sequence.

TABLE 1
Examples of Selectable Genes and their Selection Agents

Selection Agent	Selectable Gene
Puromycin	Puromycin-N-acetyltransferase
Methotrexate	Dihydrofolate reductase
Cadmium	Metallothionein
PALA	CAD
Xyl-A-or adenosine and 2'-deoxycoformycin	Adenosine deaminase
Adenine, azaserine, and coformycin	Adenylate deaminase
6-Azauridine, pyrazofuran	UMP Synthetase
Mycophenolic acid	IMP 5'-dehydrogenase
Mycophenolic acid with limiting xanthine	Xanthine-guanine phosphoribosyltransferase
Hypoxanthine, aminopterin, and thymidine (HAT)	Mutant HGPRTase or mutant thymidine kinase
5-Fluorodeoxyuridine	Thymidylate synthetase
Multiple drugs <i>e.g.</i> adriamycin, vincristine or colchicine	P-glycoprotein 170
Aphidicolin	Ribonucleotide reductase
Methionine sulfoximine	Glutamine synthetase

β -Aspartyl hydroxamate or Albizzin	Asparagine synthetase
Canavanine	Arginosuccinate synthetase
α -Difluoromethylornithine	Ornithine decarboxylase
Compactin	HMG-CoA reductase
Tunicamycin	<i>N</i> -Acetylglucosaminyl transferase
Borrelidin	Threonyl-tRNA synthetase
Ouabain	$\text{Na}^+ \text{K}^+$ -ATPase

The preferred selectable genes are amplifiable genes. As used herein, the term "amplifiable gene" refers to a gene which is amplified (*i.e.* additional copies of the gene are generated which survive in intrachromosomal or extrachromosomal form) under certain conditions. The amplifiable gene(s) usually encodes an enzyme (*i.e.* an amplifiable marker) which is required for growth of eukaryotic cells under those conditions. For example, the gene may encode DHFR which is amplified when a host cell transformed therewith is grown in Mtx. According to Kaufman, the selectable genes in Table 1 above can also be considered amplifiable genes. An example of a selectable gene which is generally not considered to be an amplifiable gene is the neomycin resistance gene (*Cepko et al., supra*).

As used herein, "selective medium" refers to nutrient solution used for growing eukaryotic cells which have the selectable gene(s) and therefore is deficient in components supplied by the selectable gene or includes a "selection agent". Commercially available media based on formulations such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are exemplary nutrient solutions. In addition, any of the media described in Ham and Wallace, *Meth. Enz.*, 58:44 (1979), Barnes and Sato, *Anal. Biochem.*, 102:255 (1980), U.S. Patent Nos. 4,767,704; 4,657,866; 4,927,762; or 4,560,655; WO 90/03430; WO 87/00195; U.S. Patent Re. 30,985; or U.S. Patent No. 5,122,469, the disclosures of all of which are incorporated herein by reference, may be used as culture media. Any of these media may be supplemented as necessary with hormones and/or other

growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics (such as GentamycinTM drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The preferred nutrient solution comprises fetal bovine serum.

The term "selection agent" refers to a substance that interferes with the growth or survival of a host cell possibly because the cell is deficient in a particular selectable gene. Examples of selection agents are presented in Table 1 above. The selection agent preferably comprises an "amplifying agent" which is defined for purposes herein as an agent for amplifying copies of the amplifiable gene or causing integration of multiple copies of the amplifiable gene into the genome, such as Mtx if the amplifiable gene is DHFR. See Table 1 for examples of amplifying agents.

As used herein, the terms "direct selection" or "direct culturing" means the first exposure to selective conditions either without MTX or GHT or with MTX, and production of a heterologous polypeptide in an amount of about 250mg/l, 400mg/l, 600mg/l or 800mg/l up to about 1g/l or more.

As used herein, the term "transcriptional initiation site" refers to the nucleic acid in the DNA construct corresponding to the first nucleic acid incorporated into the primary transcript, *i.e.*, the mRNA precursor, which site is generally provided at, or adjacent to, the 5' end of the DNA construct.

The term "transcriptional termination site" refers to a sequence of DNA, normally represented at the 3' end of the DNA construct, that causes RNA polymerase to terminate transcription.

As used herein, "transcriptional regulatory region" refers to a region of the DNA construct that regulates transcription of the selectable gene and the product gene. The transcriptional regulatory region normally refers to a promoter sequence (*i.e.* a region of DNA involved in binding of RNA polymerase to initiate transcription) which can be constitutive or inducible and, optionally, an enhancer (*i.e.* a *cis*-acting DNA element, usually from about 10-300 bp, that acts on a promoter to increase its transcription).

As used herein, "product gene" refers to DNA that encodes a desired protein or polypeptide product. Any product gene that is capable of expression in a host cell may be used, although the methods of the invention are particularly suited for obtaining high-level expression of a product gene that is not also a selectable or amplifiable gene. Accordingly, the protein or polypeptide encoded by a product gene typically will be one that is not necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. For example, product genes suitably encode a peptide, or may encode a polypeptide sequence of amino acids for which the chain length is sufficient to produce higher levels of tertiary and/or quaternary structure.

Examples of bacterial polypeptides or proteins include, e.g., alkaline phosphatase and β -lactamase. Examples of mammalian polypeptides or proteins include molecules such as renin; a growth hormone, including human growth hormone, and bovine growth hormone; growth hormone releasing factor; parathyroid hormone; thyroid stimulating hormone; lipoproteins; alpha-1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; follicle stimulating hormone; calcitonin; luteinizing hormone; glucagon; clotting factors such as factor VIIIC, factor IX, tissue factor, and von Willebrands factor; anti-clotting factors such as Protein C; atrial natriuretic factor; lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; RANTES (regulated on activation normally T-cell expressed and secreted); human macrophage inflammatory protein (MIP-1-alpha); a serum albumin such as human serum albumin; mullerian-inhibiting substance; relaxin A-chain; relaxin B-chain; prorelaxin; mouse gonadotropin-associated peptide; a microbial protein, such as beta-lactamase; DNase; inhibin; activin; vascular endothelial growth factor (VEGF); receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as bone-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), or a nerve growth factor such as NGF- β ; platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF and bFGF; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta; including TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, or TGF- β 5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-I), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), e.g., M-CSF, GM-CSF,

and G-CSF; interleukins (ILs), *e.g.*, IL-1 to IL-10; superoxide dismutase; T-cell receptors; surface membrane proteins; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; regulatory proteins; antibodies; chimeric proteins such as immunoadhesins and fragments of any of the above-listed polypeptides.

The product gene preferably does not consist of an anti-sense sequence for inhibiting the expression of a gene present in the host. Preferred proteins herein are therapeutic proteins such as TGF- β , TGF- α , PDGF, EGF, FGF, IGF-I, DNase, plasminogen activators such as t-PA, clotting factors such as tissue factor and factor VIII, hormones such as relaxin and insulin, cytokines such as IFN- γ , chimeric proteins such as TNF receptor IgG immunoadhesin (TNFr-IgG) or antibodies such as anti-IgE. An example of an antibody that can be produced with the pSV.IDP plasmid (Figure 4) is anti-HER2 Neu antibody, 2C4, as provided in Example 1, *supra*.

The term "intron" as used herein refers to a nucleotide sequence present within the transcribed region of a gene or within a messenger RNA precursor, which nucleotide sequence is capable of being excised, or spliced, from the messenger RNA precursor by a host cell prior to translation. Introns suitable for use in the present invention are suitably prepared by any of several methods that are well known in the art, such as purification from a naturally occurring nucleic acid or *de novo* synthesis. The introns present in many naturally occurring eukaryotic genes have been identified and characterized. Mount, Nuc. Acids Res., 10:459 (1982). Artificial introns comprising functional splice sites also have been described. Winey *et al.*, Mol. Cell Biol., 9:329 (1989); Gatermann *et al.*, Mol. Cell Biol., 9:1526 (1989). Introns may be obtained from naturally occurring nucleic acids, for example, by digestion of a naturally occurring nucleic acid with a suitable restriction endonuclease, or by PCR cloning using primers complementary to sequences at the 5' and 3' ends of the intron. Alternatively, introns of defined sequence and length may be prepared synthetically using various methods in organic chemistry. Narang *et al.*, Meth. Enzymol., 68:90 (1979); Caruthers *et al.*, Meth. Enzymol., 154:287 (1985); Froehler *et al.*, Nuc. Acids Res., 14:5399 (1986).

As used herein "splice donor site" or "SD" refers to the DNA sequence immediately surrounding the exon-intron boundary at the 5' end of the intron, where the "exon" comprises the nucleic acid 5' to the intron. Many splice donor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., 195:247-259 (1987) provides a review of these. An "efficient splice donor sequence" refers to a nucleic acid sequence encoding a splice donor site wherein the efficiency of splicing of

messenger RNA precursors having the splice donor sequence is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. Examples of efficient splice donor sequences include the wild type (WT) ras splice donor sequence and the GAC:GTAAGT sequence of Example 3. Other efficient splice donor sequences can be readily selected using the techniques for measuring the efficiency of splicing disclosed herein.

The terms "PCR" and "polymerase chain reaction" as used herein refer to the *in vitro* amplification method described in US Patent No. 4,683,195 (issued July 28, 1987). In general, the PCR method involves repeated cycles of primer extension synthesis, using two DNA primers capable of hybridizing preferentially to a template nucleic acid comprising the nucleotide sequence to be amplified. The PCR method can be used to clone specific DNA sequences from total genomic DNA, cDNA transcribed from cellular RNA, viral or plasmid DNAs. Wang & Mark, in PCR Protocols, pp.70-75 (Academic Press, 1990); Scharf, in PCR Protocols, pp. 84-98; Kawasaki & Wang, in PCR Technology, pp. 89-97 (Stockton Press, 1989). Reverse transcription-polymerase chain reaction (RT-PCR) can be used to analyze RNA samples containing mixtures of spliced and unspliced mRNA transcripts. Fluorescently tagged primers designed to span the intron are used to amplify both spliced and unspliced targets. The resultant amplification products are then separated by gel electrophoresis and quantitated by measuring the fluorescent emission of the appropriate band(s). A comparison is made to determine the amount of spliced and unspliced transcripts present in the RNA sample.

One preferred splice donor sequence is a "consensus splice donor sequence". The nucleotide sequences surrounding intron splice sites, which sequences are evolutionarily highly conserved, are referred to as "consensus splice donor sequences". In the mRNAs of higher eukaryotes, the 5' splice site occurs within the consensus sequence AG:GUAAGU (wherein the colon denotes the site of cleavage and ligation). In the mRNAs of yeast, the 5' splice site is bounded by the consensus sequence :GUAUGU. Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986).

The expression "splice acceptor site" or "SA" refers to the sequence immediately surrounding the intron-exon boundary at the 3' end of the intron, where the "exon" comprises the nucleic acid 3' to the intron. Many splice acceptor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., 195:247-259 (1987) provides a review of these. The preferred splice acceptor site is an efficient splice acceptor site which refers to a nucleic acid sequence encoding a splice

acceptor site wherein the efficiency of splicing of messenger RNA precursors having the splice acceptor site is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. The splice acceptor site may comprise a consensus sequence. In the mRNAs of higher eukaryotes, the 3' splice acceptor site occurs within the consensus sequence (U/C)₁₁NCAG:G. In the mRNAs of yeast, the 3' acceptor splice site is bounded by the consensus sequence (C/U)AG:. Padgett, *et al.*, *supra*.

As used herein "culturing for sufficient time to allow amplification to occur" refers to the act of physically culturing the eukaryotic host cells which have been transformed with the DNA construct in cell culture media containing the amplifying agent, until the copy number of the amplifiable gene (and preferably also the copy number of the product gene) in the host cells has increased relative to the transformed cells prior to this culturing.

The term "expression" as used herein refers to transcription or translation occurring within a host cell. The level of expression of a product gene in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell or the amount of the protein encoded by the product gene that is produced by the cell. For example, mRNA transcribed from a product gene is desirably quantitated by northern hybridization or quantitative real-time PCR. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 7.3-7.57 (Cold Spring Harbor Laboratory Press, 1989). Protein encoded by a product gene can be quantitated either by assaying for the biological activity of the protein or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay using antibodies that are capable of reacting with the protein. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 18.1-18.88 (Cold Spring Harbor Laboratory Press, 1989).

Modes for Carrying Out the Invention

Methods and compositions are provided for enhancing the stability and/or copy number of a transcribed sequence in order to allow for elevated levels of a RNA sequence of interest. In general, the methods of the present invention involve transfecting a eukaryotic host cell with an expression vector comprising both a product gene encoding a desired polypeptide and fused selectable genes.

Selectable genes and product genes may be obtained from genomic DNA, cDNA transcribed from cellular RNA, or by *in vitro* synthesis. For example, libraries are screened with

probes (such as antibodies or oligonucleotides of about 20-80 bases) designed to identify the selectable gene or the product gene (or the protein(s) encoded thereby). Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures as described in chapters 10-12 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the selectable gene or product gene is to use PCR methodology as described in section 14 of Sambrook *et al.*, *supra*.

A preferred method of practicing this invention is to use carefully selected oligonucleotide sequences to screen cDNA libraries from various tissues known to contain the selectable gene or product gene. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized.

The oligonucleotide generally is labeled such that it can be detected upon hybridization to DNA in the library being screened. The preferred method of labeling is to use ³²P- labeled ATP with polynucleotide kinase, as is well known in the art, to radiolabel the oligonucleotide. However, other methods may be used to label the oligonucleotide, including, but not limited to, biotinylation or enzyme labeling.

Sometimes, the DNA encoding the fused selectable genes and product gene is preceded by DNA encoding a signal sequence having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the expression vector, or it may be a part of the selectable gene or product gene that is inserted into the expression vector. If a heterologous signal sequence is used, it preferably is one that is recognized and processed (*i.e.*, cleaved by a signal peptidase) by the host cell. For yeast secretion the native signal sequence may be substituted by, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Pat. No. 5,010,182 issued 23 April 1991), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression the native signal sequence of the protein of interest is satisfactory, although other mammalian signal sequences may be suitable, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders, for example, the herpes simplex gD signal. The DNA for such precursor region is ligated in reading frame to the fused selectable genes or product gene.

As shown in Figure 1, the fused selectable genes are generally provided at the 5' end of the DNA construct and are followed by the product gene (which would be inserted into the linker site). Therefore, the full-length (non-spiced) message will contain, for example, the PURO-DHFR fusion as the first open reading frame and will therefore generate PURO-DHFR protein to allow selection of stable transfectants. The full length message is not expected to generate appreciable amounts of the protein of interest as the second AUG in a dicistronic message is an inefficient initiator of translation in mammalian cells (Kozak, J. Cell Biol., **115**: 887-903 (1991)).

The fused selectable genes are positioned within an intron. Introns are noncoding nucleotide sequences, normally present within many eukaryotic genes, which are removed from newly transcribed mRNA precursors in a multiple-step process collectively referred to as splicing.

A single mechanism is thought to be responsible for the splicing of mRNA precursors in mammalian, plant, and yeast cells. In general, the process of splicing requires that the 5' and 3' ends of the intron be correctly cleaved and the resulting ends of the mRNA be accurately joined, such that a mature mRNA having the proper reading frame for protein synthesis is produced. Analysis of a variety of naturally occurring and synthetically constructed mutant genes has shown that nucleotide changes at many of the positions within the consensus sequences at the 5' and 3' splice sites have the effect of reducing or abolishing the synthesis of mature mRNA. Sharp, Science, **235**:766 (1987); Padgett, *et al.*, Ann. Rev. Biochem., **55**:1119 (1986); Green, Ann. Rev. Genet., **20**:671 (1986). Mutational studies also have shown that RNA secondary structures involving splicing sites can affect the efficiency of splicing. Solnick, Cell, **43**:667 (1985); Konarska, *et al.*, Cell, **42**:165 (1985).

The length of the intron may also affect the efficiency of splicing. By making deletion mutations of different sizes within the large intron of the rabbit beta-globin gene, Wieringa, *et al.* determined that the minimum intron length necessary for correct splicing is about 69 nucleotides. Cell, **37**:915 (1984). Similar studies of the intron of the adenovirus E1A region have shown that an intron length of about 78 nucleotides allows correct splicing to occur, but at reduced efficiency. Increasing the length of the intron to 91 nucleotides restores normal splicing efficiency, whereas truncating the intron to 63 nucleotides abolishes correct splicing. Ulfendahl, *et al.*, Nuc. Acids Res., **13**:6299 (1985).

To be useful in the invention, the intron must have a length such that splicing of the intron from the mRNA is efficient. The preparation of introns of differing lengths is a routine matter,

involving methods well known in the art, such as *de novo* synthesis or *in vitro* deletion mutagenesis of an existing intron. Typically, the intron will have a length of at least about 150 nucleotides, since introns which are shorter than this tend to be spliced less efficiently. The upper limit for the length of the intron can be up to 30 kB or more. However, as a general proposition, the intron is generally less than about 10 kB in length.

The intron is modified to contain the fused selectable genes not normally present within the intron using any of the various known methods for modifying a nucleic acid *in vitro*. Typically, the fused selectable genes will be introduced into an intron by first cleaving the intron with a restriction endonuclease, and then covalently joining the resulting restriction fragments to the fused selectable genes in the correct orientation for host cell expression, for example by ligation with a DNA ligase enzyme.

The DNA construct is dicistronic, *i.e.* the fused selectable genes and product gene are both under the transcriptional control of a single transcriptional regulatory region. As mentioned above, the transcriptional regulatory region comprises a promoter. Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, J. Biol. Chem., **255**:2073 (1980)) or other glycolytic enzymes (Hess *et al.*, J. Adv. Enzyme Reg., **7**:149 (1968); and Holland, Biochemistry, **17**:4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in Hitzeman *et al.*, EP 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Expression control sequences are known for eukaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide.

Product gene transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40) or cytomegalovirus (CMV), from heterologous mammalian promoters, e.g. the actin promoter or an immunoglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the product gene, provided such promoters are compatible with the host cell systems. Promoters endogenous to the host cell system, such as the CHO Elongation Factor 1 alpha promoter may also be used.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers *et al.*, Nature, 273:113 (1978); Mulligan and Berg, Science, 209:1422-1427 (1980); Pavlakis *et al.*, Proc. Natl. Acad. Sci. USA, 78:7398-7402 (1981). The immediate early promoter of the human cytomegalovirus (CMV) is conveniently obtained as a *Hind*III E restriction fragment. Greenaway *et al.*, Gene, 18:355-360 (1982). A system for expressing DNA in mammalian hosts using the bovine papilloma virus as a vector is disclosed in U.S. 4,419,446. A modification of this system is described in U.S. 4,601,978. See also Gray *et al.*, Nature, 295:503-508 (1982) on expressing cDNA encoding immune interferon in monkey cells; Reyes *et al.*, Nature, 297:598-601 (1982) on expression of human β -interferon cDNA in mouse cells under the control of a thymidine kinase promoter from herpes simplex virus, Canaani and Berg, Proc. Natl. Acad. Sci. USA, 79:5166-5170 (1982) on expression of the human interferon β 1 gene in cultured mouse and rabbit cells, and Gorman *et al.*, Proc. Natl. Acad. Sci. USA, 79:6777-6781 (1982) on expression of bacterial CAT sequences in CV-1 monkey kidney cells, chicken embryo fibroblasts, Chinese hamster ovary cells, HeLa cells, and mouse NIH-3T3 cells using the Rous sarcoma virus long terminal repeat as a promoter.

Preferably the transcriptional regulatory region in higher eukaryotes comprises an enhancer sequence. Enhancers are relatively orientation and position independent having been found 5' (Lainins *et al.*, Proc. Natl. Acad. Sci. USA, 78:993 (1981)) and 3' (Lusky *et al.*, Mol. Cell Bio., 3:1108 (1983)) to the transcription unit, within an intron (Banerji *et al.*, Cell, 33:729 (1983)) as well as within the coding sequence itself (Osborne *et al.*, Mol. Cell Bio., 4:1293 (1984)). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -

fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer (CMV), the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yaniv, *Nature*, 297:17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the product gene, but is preferably located at a site 5' from the promoter.

The DNA construct of the present invention has a transcriptional initiation site following the transcriptional regulatory region and a transcriptional termination region following the product gene (see, e.g., Figure 1). These sequences are provided in the DNA construct using techniques which are well known in the art.

The DNA construct normally forms part of an expression vector which may have other components such as an origin of replication (*i.e.*, a nucleic acid sequence that enables the vector to replicate in one or more selected host cells) and, if desired, one or more additional selectable gene(s). Construction of suitable vectors containing the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

Generally, in cloning vectors the origin of replication is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known. The 2 μ plasmid origin of replication is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (the SV40 origin may typically be used only because it contains the early promoter).

Most expression vectors are "shuttle" vectors, *i.e.*, they are capable of replication in at least one class of organisms but can be transfected into another organism for expression. For example, a vector is cloned in *E. coli* and then the same vector is transfected into yeast or mammalian cells for expression even though it is not capable of replicating independently of the host cell chromosome.

For analysis to confirm correct sequences in plasmids constructed, plasmids from the transformants are prepared, analyzed by restriction, and/or sequenced by the method of Messing *et*

al., Nucleic Acids Res., 9:309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology, 65:499 (1980).

The expression vector having the DNA construct prepared as discussed above is transformed into a eukaryotic host cell. Suitable host cells for cloning or expressing the vectors herein are yeast or higher eukaryote cells.

Eukaryotic microbes such as filamentous fungi or yeast are suitable hosts for vectors containing the product gene. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used among lower eukaryotic host microorganisms. However, a number of other genera, species, and strains are commonly available and useful herein, such as *S. pombe* (Beach and Nurse, Nature, 290:140 (1981)), *Kluyveromyces lactis* (Louvcourt *et al.*, J. Bacteriol., 737 (1983)), *kyarrowia* (EP 402,226), *Pichia pastoris* (EP 183,070), *Trichoderma reesia* (EP 244,234), *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 (1979)), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 (1983); Tilburn *et al.*, Gene, 26:205-221 (1983); Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 (1985)).

Suitable host cells for the expression of the product gene are derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* host cells have been identified. See, e.g., Luckow *et al.*, Bio/Technology, 6:47-55 (1988); Miller *et al.*, in Genetic Engineering, Setlow, J.K. *et al.*, eds., Vol. 8 (Plenum Publishing, 1986), pp. 277-279; and Maeda *et al.*, Nature, 315:592-594 (1985). A variety of such viral strains are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

Plant cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco can be utilized as hosts. Typically, plant cells are transfected by incubation with certain strains of the bacterium *Agrobacterium tumefaciens*, which has been previously manipulated to contain the product gene. During incubation of the plant cell culture with *A. tumefaciens*, the product gene is

transferred to the plant cell host such that it is transfected, and will, under appropriate conditions, express the product gene. In addition, regulatory and signal sequences compatible with plant cells are available, such as the nopaline synthase promoter and polyadenylation signal sequences. Depicker *et al.*, J. Mol. Appl. Gen., 1:561 (1982). In addition, DNA segments isolated from the upstream region of the T-DNA 780 gene are capable of activating or increasing transcription levels of plant-expressible genes in recombinant DNA-containing plant tissue. EP 321,196 published 21 June 1989.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years (Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)). Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen Virol., 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); dp12.CHO cells (EP 307,247 published 15 March 1989); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, Annals N.Y. Acad. Sci., 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) may be used. General aspects of mammalian cell host system transformations have been described by Axel in U.S. 4,399,216 issued 16 August 1983. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829

(1979). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

In preferred embodiments the DNA is introduced into the host cells using electroporation, lipofection or polyfection techniques. In a particularly preferred embodiment, the transfection is performed in a spinner vessel as illustrated by Example 3 or in some other form of suspension culture. Transfection performed in a spinner vessel is also referred to as "spinner transfection". Culturing the cells in suspension allows them to reach a cell density of at least about 5×10^5 /ml and more preferably at least about 1.5×10^6 /ml prior to transfection. See Andreason, J. Tiss. Cult. Meth., 15:56-62 (1993), for a review of electroporation techniques useful for practicing the claimed invention. It was discovered that these techniques for introducing the DNA construct into the host cells are preferable over calcium phosphate precipitation techniques insofar as the latter could cause the DNA to break up and form concatemers.

The mammalian host cells used to express the product gene herein may be cultured in a variety of media as discussed in the definitions section above. The media is formulated to provide selective nutrient conditions or a selection agent to select transformed host cells which have taken up the DNA construct (either as an intra- or extra-chromosomal element). To achieve selection of the transformed eukaryotic cells, the host cells may be grown in cell culture plates and individual colonies expressing one or both of the selectable genes (and thus the product gene) can be isolated and grown in growth medium under defined conditions. The host cells are then analyzed for transcription and/or transformation as discussed below. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA or mRNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly ^{32}P . However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescence, enzymes, or the like. Alternatively, antibodies may be employed that can recognize specific

duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. With immunohistochemical staining techniques, a cell sample is prepared, typically by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product coupled, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like. A particularly sensitive staining technique suitable for use in the present invention is described by Hsu *et al.*, Am. J. Clin. Path., 75:734-738 (1980).

In the preferred embodiment protein expression is measured using ELISA as described in Example 1 herein.

The product of interest preferably is recovered from the culture medium as a secreted polypeptide, although it also may be recovered from host cell lysates when directly expressed without a secretory signal. When the product gene is expressed in a recombinant cell other than one of human origin, the product of interest is completely free of proteins or polypeptides of human origin. However, it is necessary to purify the product of interest from recombinant cell proteins or polypeptides to obtain preparations that are substantially homogeneous as to the product of interest. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The product of interest thereafter is purified from contaminant soluble proteins and polypeptides, for example, by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel electrophoresis using, for example, Sephadex G-75; chromatography on plasminogen columns to bind the product of interest and protein A Sepharose columns to remove contaminants such as IgG.

The following examples are offered by way of illustration only and are not intended to limit the invention in any manner. All patent and literature references cited herein are expressly incorporated by reference.

EXAMPLE 1

2C4 production using the fusion construct expression vector

Vectors related to those described by Lucas *et al* (Lucas BK, Giere LM, DeMarco RA, Shen A, Chisholm V and Crowley C. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. (1996) Nucleic Acids Res. 24(9), 1774-1779.), which contain an intron between the SV40 promoter and enhancer and the cDNA that encodes the polypeptide of interest, were constructed. The intron is bordered on its 3' and 5' ends, respectively, by a splice donor site derived from cytomegalovirus immediate early gene (CMVIE), and a splice acceptor site from an IgG heavy chain variable region (V_H) gene (Eaton *et al.*, *Biochem.*, 25:8343 (1986)). The splice sites selected provide slightly inefficient splicing such that only about 90% of the transcripts produced are intron free. Previous studies have demonstrated that when a selectable marker such as DHFR is integrated within this intron, as in the plasmid pSV.ID, marker gene transcription proceeds from any unspliced transcripts, providing a highly efficient means of maintaining linkage between the expression of the marker gene and the cDNA of interest as well as enhanced product expression relative to expression of the marker gene.

Vectors containing a murine puromycin/DHFR fusion sequence in the intron following the SV40 promoter elements were constructed by linearizing a pSV.IPUR plasmid, which contained the puromycin resistance gene in an intron following the SV40 promoter/enhancer (pSV.IPUR, Figures 1 and 2), with Hpa I immediately following the end of the puromycin ORF. A 564 bp PCR fragment containing the entire coding region for the murine DHFR gene was subsequently ligated into this linearized vector 3' of the puromycin resistance gene. The stop codon TAG between the puromycin resistance gene and the DHFR gene was deleted by site-directed mutagenesis resulting in a pSV.I plasmid containing a Puro/DHFR fusion gene within the intron of the expression cassette (pSV.IPD, Figures 1 and 4).

The cDNA of the Heavy chain (HC) and light chain (LC) sequences of an anti-HER2 Neu antibody, 2C4, were inserted into pSV.IPD as shown in Figure 6. The sequence of the resulting pSV.IPD.2C4 vector is shown in Figure 7. Data collected using the pSV.IPD.2C4 vector are shown in Table 2.

Additionally, a vector containing only a murine DHFR sequence within the intron (pSV.ID) was prepared. The DNA sequence for the pSV.ID vector is shown in Figure 3. The preparation of such vectors is disclosed in U.S. Patent No. 5,561,053, which is herein incorporated by reference. Into that vector, the HC and LC sequences of monoclonal antibodies to VEGF were inserted. The sequence of the resulting pSV.ID.VEGF vector is shown in Figure 5.

Plasmid DNA's that contained either the Puro/DHFR fusion sequences in the intron or murine DHFR alone preceding cDNA sequences for HC and LC of 2C4 and anti-VEGF, respectively were introduced into CHO DHFR minus cells by lipofection. Briefly, for transfection, 4 million CHO DUX-B11 (DHFR minus) were seeded in 10 cm plates the day before transfection. On the day of transfection, 4 ug DNA was mixed with 300 ul of serum free medium and 25 ul of polyfect from Qiagen. The mixture was incubated at room temperature for 5 to 10 minutes and added to the cells. Cells were fed with fresh glycine, hypoxanthine and thymidine-free (GHT-free) medium and twenty-four hours later, were trypsinized and selected in fresh GHT- free medium with 0 – 5 nM of methotrexate (MTX) in order to select for stable DHFR+ clones. Approximately 300 – 400 individual clones were selected in this first round of screening for measurement of protein expression levels. Clones from each vector which expressed the highest levels of antibody were then re-exposed to higher levels of methotrexate to affect a second round of gene amplification and selection. The screening process was repeated on all available clones, the highest of which were exposed to a third round of amplification. The methotrexate concentrations used during amplification using the pSV.ID-derived vector was 50 to 1000 nM in the 2nd round and 200 to 1000 nM in the 3rd round. These concentrations are typically required to achieve growth-limiting toxicity, which is required to achieve sufficient selective pressure for gene amplification. Concentrations required to reach this same degree of toxicity using the pSV.ID-derived vectors were remarkably lower.

The level of antibody expression was determined by seeding cells in 1 ml of serum-free F12:DMEM-based media supplemented with protein hydrolysate and amino acids in 24 well dishes at 3 X 10⁶ cells/ml or in 100 ul of similar media in individual wells of a 96 well plate. Growth media was collected after 3-4 days and titers were assayed by an ELISA directed towards the intact IgG molecule. In experiments where cells were not seeded at equal cell densities, a fluorescent measure of viable cell number was performed on each well in order to normalize expression data. An Intact IgG ELISA was performed on microtiter plates which used a capture

antisera directed to framework Fab residues common in both antibodies. Media samples were added to the wells followed by washing and a horseradish peroxidase labeled second antibody directed towards common framework Fc residues was used for detection.

Table 2 presents expression level distributions of clones isolated during each round of screening of anti VEGF clones, which resulted from transfection with the plasmid containing only the DHFR sequence in the intron (pSV.ID.aVEGF), and 2C4 clones that were created using the Puro/DHFR fusion sequence in the same intron (pSV.IDP.2C4). The distribution of expression levels seen in the case of anti VEGF is typical of the performance of the vector containing only the murine DHFR gene in the intron (pSV.ID). All isolates identified in the first and second rounds of screening have relatively low expression levels. In the intial selection round, no clones with expression above 5 were isolated. At least three rounds of amplification are required to identify clones capable of specific productivity greater than 50. The 2C4 clones were screened after the first exposure to methotrexate (0-2.5 nM) and the most productive of these were exposed to a second round of amplification in 10-25 nM MTX. Cells surviving this amplification were pooled and exposed to 3rd round amplification prior to selection for further screening. In contrast to the pSV.ID vtor, using the pSV.IDP vector, clones with an expression level of up to 25 were identified even in the first round of screening. Clones with an expression level greater than 25 represented 95% of the population after their third round of amplification and screening.

The data from Example 1 indicates that use of the Puro/DHFR fusion protein as the selectable marker allows for faster, more efficient isolation of highly productive CHO clones using significantly lower levels of methotrexate. The data suggests that exposure to low concentrations and stepwise increments in methotrexate allow for the efficient initial selection of highly expressing clones and subsequent gene amplification. Exposure to excessively high concentrations of methotrexate or large incremental increases in exposure often does not yield increases in gene expression since cells rapidly acquire methotrexate resistance through non-gene amplification mechanisms. Importantly, the data also shows that the Puro/DHFR fusion protein provides an unexpectedly impaired activity of the DHFR gene product or an enhanced sensitivity to methotrexate, which results in a highly stringent initial selection step, and allows efficient gene amplification at concentrations of methotrexate not frequently associated with the acquisition of drug resistance through alternative mechanisms. The ability to select cells which have incorporated the plasmid either in the presence of puromycin or methotrexate, prior to initiating exposure to

methotrexate also provides a means of transferring this efficient system to DHFR (positive) host cells.

For Example 1 the structure of the expressed antibody has been extensively characterized. The proteins generated from the pSV.IPD are indistinguishable from the antibody produced by the pSV.ID vector, with no apparent increase of free heavy or light chain expressed by the pool.

TABLE 2. PERCENTAGES OF pSV.IPD.2C4 CLONES ISOLATED AT VARIOUS EXPRESSION LEVELS AFTER MTX EXPOSURE¹

Expression Level ²	pSV.ID.aVEGF 1st Rd	pSV.IPD.2C4 1st Rd	pSV.ID.aVEGF 3rd Rd	pSV.IPD.2C4 3rd Rd
<1	71	16	0	0
1-5	29	67	0	0
5-10	0	14	2	3
10-25	0	3	15	4
25-50	0	0	35	21
50-100	0	0	46	61
100-150	0	0	2	3

¹MTX concentration for Control SD vector = 0-10 nM 1st round, 50 –1000 nM 2nd round, 200-1000 nM, 3rd round. SD- Puro/DHFR vector = 2.5 nM 1st round, 25 nM 2nd round, 100 nM 3rd round.

² Expression levels are in mg/ml or (mg/ml)/Fluorescent Unit

This example demonstrate the general applicability of the Puro/DHFR fusion sequence for selection of highly productive recombinant cell lines following minimal exposure to MTX.

EXAMPLE 2

Recombinant protein production using a pSV.I construct containing DHFR and a fusion gene other than Puro

Constructs can also be produced that contain a fusion sequence of an alternative selectable marker and DHFR within an intron region as described in Example 1. For instance

starting with the vector pSVID, the coding sequences for the neomycin resistance gene (Neo), hygromycin resistance gene (Hygro), glutamine synthase (GS), thymidine kinase (TK) or zeocin (Zeo) could be inserted in frame with the start site of the murine DHFR sequence contained within the intron. The stop codon of this inserted gene would then be removed using site directed mutagenesis according to example 1. Depending upon the phenotype of the host cell selected, cells incorporating the plasmid could then be selected using either GHT-free or MTX containing media as described in examples 1-3 or using an appropriate quantity of the alternative selective agent. Gene expression by the resulting clones could then be amplified in the presence of increased levels of methotrexate.

EXAMPLE 3

Direct Selection with plasmids SV.IPD.HP and CMV.IPD.HP after spinner transfection

DP12 CHO cells were grown in growth medium with 5% FBS (fetal bovine serum) and 1X GHT (glycine, hypoxanthine and thymidine). The process typically took about 4 days. On day 1, cells were seeded at 4×10^5 /ml in 400 ml growth medium in a 500 ml spinner vessel and grown for 2 days at 37°C . On day 3, the exponentially grown cells were seeded at 1.5×10^6 cells/ml in a 250 ml spinner vessel containing 200 ml of growth medium plus 5% FBS and 1X GHT. The cells were grown for 1 to 2 hours at 37°C before transfection. During that time, serum-free growth medium and 1X GHT was warmed to 37°C . 400 μg plasmid construct DNA and 1 ml of Lipofectamine 2000[®] (Qiagen) were separately diluted into 25 ml of warm serum-free medium in 50 ml Falcon tubes. The solutions in the tubes were combined and incubated at room temperature for 30 minutes. The cells were then transfected with plasmid constructs pSV.IPD.HP and pCMV.IPD.HP, which constructs are illustrated in Figures 13 and 14, respectively. At the end of incubation, the cells were transfected by adding all 50 ml of the mixture of diluted plasmid construct and Lipofectamine 2000[®] to the 250 ml spinner vessel containing cells in serum-free medium, and the cells continued to grow at 37°C for about 24 hours. On day 4, 250 ml of transfected cells were centrifuged at 1000 rpm for 5 minutes to collect the pellet. The transfection efficiency was monitored by transfecting cells with a GFP plasmid followed by FACS analysis 24 hours after transfection. The transfection efficiency with this protocol was typically approximately 55 to 70 % in CHO cells as shown in Figure 8.

After the transfection, cells were centrifuged to collect the pellet. The pellet was then resuspended in growth medium containing methotrexate (MTX) ranging from 10 to 100 nM for either SV40 or CMV based constructs. Approximately 100 clones survived the direct selection. Cell growth medium was changed every 3 to 4 days. At approximately 2 weeks after transfection, individual clones were picked and grown in 96-well plates in growth medium containing MTX. Heterologous polypeptide expression levels were evaluated by ELISA. Figures 10-1, 10-2, and 11 show the results from 25 nM and 50 nM MTX selection. Figure 9 shows heterologous polypeptide expression levels of clones from a traditional 10 nM MTX selection where the cells were not transfected in a spinner flask.

It took about 1 week for cells to grow confluent in a 96-well plate. When they were confluent, the growth medium was removed and commercially available enriched cell culture medium (which includes 1x GHT but no MTX) was added into each well. On the day after adding the commercially available enriched cell culture medium, the plate was incubated at 33 °C for 5-6 days before performing an ELISA assay to quantitate the amount of humanized monoclonal antibody produced by the cells. ELISA was typically performed with serial dilutions of the commercially available enriched cell culture medium. Results from a humanized monoclonal antibody production were shown in Figures 9, 10-1, 10-2 and 11.

The four clones producing the greatest amount over 100 µg/ml of intact IgG based on direct selection at 25 nM MTX using a CMV-based construct were scaled up from a 96-well plate to a 6-well plate and then to a 10 cm plate. Cells were seeded at 3X10⁵/ml in 200 ml volume in a 250 ml spinner vessel in serum-free growth medium with 2 µg/ml human insulin and 1X Trace Elements (TE). Cells were initially passaged at either two- or three-day intervals with medium exchange. Then they were passaged at either three- or four-day intervals for about 6 weeks before bioreactor evaluation. At each passage time, cell viability and count number were monitored. To determine the cell growth after serum-free adaptation, a spinner vessel growth experiment was performed. Cells were seeded at 3X10⁵ cells/ml into 400 ml of growth medium with 2 µg/ml recombinant human insulin and 1X TE in a 500 ml spinner vessel on day 1. On each day, packed cell volume (PCV) was monitored until day 5. PCVs reached between 0.4 % to 0.6% by day 4. Two serum-free adapted clones from 25 nM MTX selection with CMV-based construct were evaluated in bioreactors. Two liter bioreactors with commercially available

enriched cell culture medium were run for a total of 14 days. The data from the titer evaluation is shown in Figure 12.

An ELISA assay of clones surviving the direct selection shows that the best clones coming out of the method described in this example produce as much product of interest as highly amplified clones from a traditional method. See Figure 16. Evaluations of 2 clones from the direct selection shows that those clones produce about 1g/L of a product of interest in a bioreactor process. Since those clones were generated from one step of a direct selection immediately after transfection, it only takes about 5 to 6 weeks to generate a stable cell line producing 1g/L of a product of interest in a bioreactor leading to significant timeline reduction, about 3 months, which is critical for efficiency of product development.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the examples presented herein, since the exemplified embodiments are intended as illustrations of certain aspects of the invention and any functionally equivalent embodiments are within the scope of this invention. The examples presented herein are not intended as limiting the scope of the claims to the specific illustrations. Indeed, various modifications of the invention, in addition to those shown and described herein and which fall within the scope of the appended claims, may become apparent to those skilled in the art from the foregoing description.

CLAIMS

What is claimed is:

1. A method of producing a host cell capable of producing a product of interest, comprising:
 - transfected a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;
 - directly culturing the transfected host cells in a selective medium;
 - allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur; and
 - selecting a host cell clone that is capable of producing at least about 250mg/l of the product of interest.
2. A method of claim 1 wherein the selective medium contains at least about 25nM methotrexate.
3. A method of claim 1 wherein the selective medium contains at least about 50nM methotrexate.
4. A method of claim 1 wherein the host cell is a CHO cell.
5. A method of claim 1 wherein the product of interest is a protein selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin; or a fragment of said protein.
6. A method of claim 1 wherein said product of interest is a humanized antibody.
7. A host cell produced according to the method of claim 1.

8. A method of producing a product of interest, comprising culturing a host cell produced according to the method of claim 1 under conditions suitable to cause expression of the product of interest in an amount at least about 250mg/l.

9. A method of claim 1 wherein the DNA construct comprises, in order 5' to 3':

- a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;
- b) a transcriptional initiation site;
- c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;
- d) a product gene encoding a product of interest; and
- e) a transcriptional termination site.

10. The method of claim 9 further comprising recovering the product of interest from the culture.

11. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a SV40 promoter.

12. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a CMV promoter.

13. A cell culture composition comprising a host cell according to claim 9 and at least about 250mg/l of the product of interest.

14. A method of producing a host cell capable of producing at least about 250mg/ml of a product of interest comprising transfecting a host cell with a DNA construct comprising in order from 5' to 3':

- a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;
- b) a transcriptional initiation site;
- c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;
- d) a product gene encoding a product of interest; and
- e) a transcriptional termination site;

wherein the transfection is performed in suspension culture.

15. A method of claim 14, wherein the DNA construct is introduced into the host cells by lipofection.

16. A method of claim 14 wherein said transfection is performed in a spinner vessel.

17. The method of claim 14 wherein the suspension culture has cell density of at least about 5×10^5 /ml at the time of transfection.

18. The method of claim 14 wherein the suspension culture has a cell density of at least about 1.5×10^5 /ml at the time of transfection

19. A method of claim 15 wherein the product of interest is selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin and a fragment of any of said product of interest.

20. A method of rapidly selecting a host cell producing a product of interest, comprising:

transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;

directly culturing the transfected host cells in a selective medium; and

allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur.

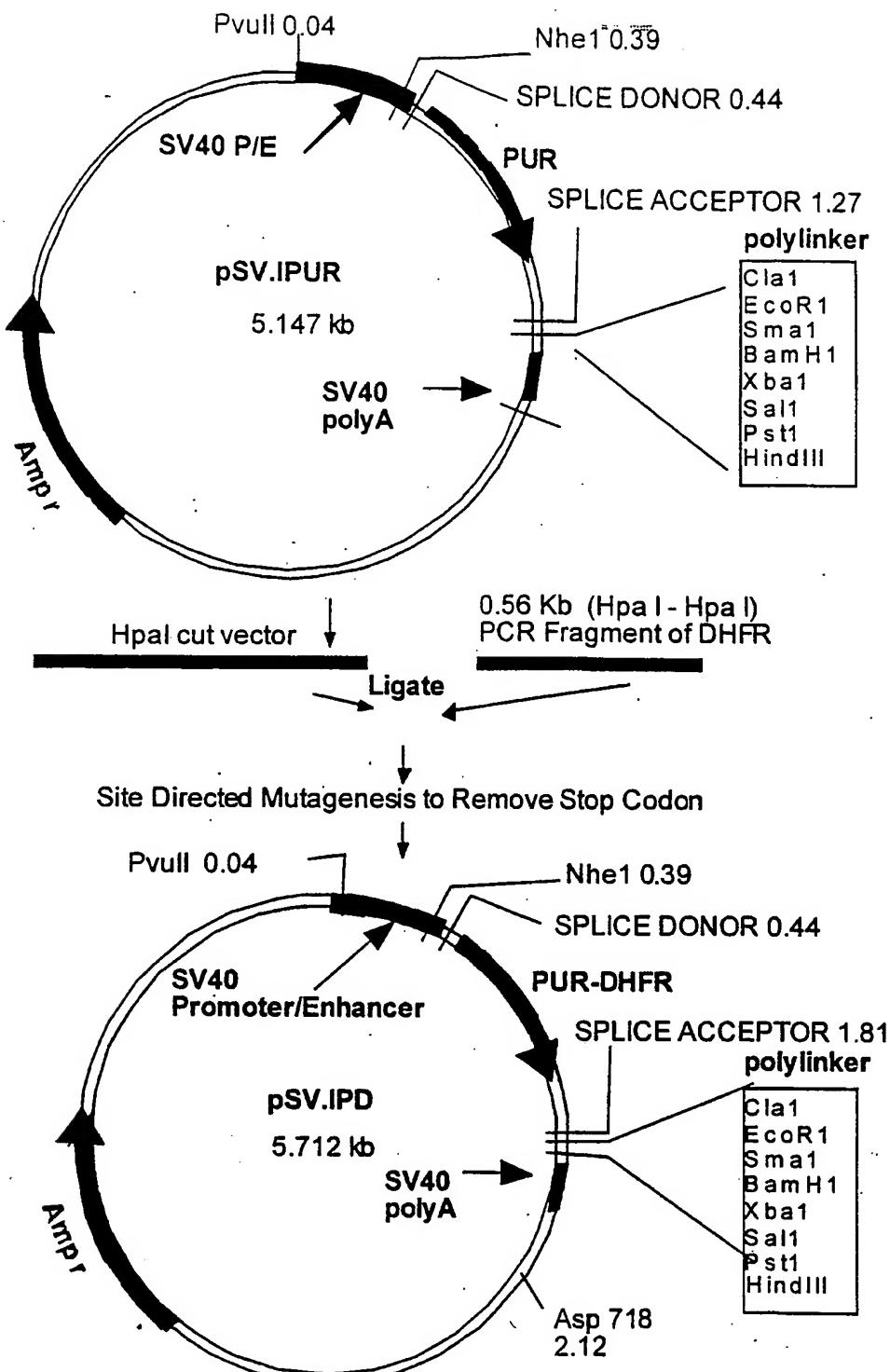


Figure 1. Construction of pSV.IPD Plasmid

Figure 2
PSV. IPUR
length: 5147 (circular)

1 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTG CAGTAGGGT CTTGGAAAGTC CCCAATCTCC CCAAGCAAGTAA
 AAGCTGAGC GGGCTGAACT TAATACTGA TCTCAGCTG CTTAGCACA
 101 GAAGATGCC AAGCATGCT CTCAAATTAGT CAGCAACCRG GTGTTGAAAG TCCCGAGGT CCCCAGGAGG CAGAAAGTATG CAAACCATGC AATCTCAAA
 CTTCAATCTGTT TTCGTAGTA CAGTTAATCA GTGTTGGTC CACACCTTC AGGGTTCGA GGGGTGTC
 201 GTCAAGRACC ATAGTCCCG CCCTRACTCC GGCATCCCG CCCCCTAACTC CGCCCAAGTC CGCCCAATTCT CGCCCACTATG GCCTGALTAT
 CAGTGTGG TAGTGGGGC GGGTGTGGC GGGTGTGGC GGGGTAGGG CGGGTAGGG CGGGTAGGA GCGGGGTAC CACTGATFA AAAAATATAA
 301 TATGGAGGG CCGAGCCGC CTCGGCCCTCT GAGCTATTC AGAAAGTATG AGGGGCTT TTGGAGGCC TAGGCTTTG CAAAGCTA GCTTATTC
 ATACCTCTCC GGCTCGGGC GAGCCGGAA CTCGATAAGG TCCTCATCAC TCCPTCGRAA ARACTCCGG ATCAGAAAAC GTTTCCTCAT CGAATAGGCC
 401 CCGGGAACGG TGCATGGRA CGGGGATTC CCGTGGCAAAG AGTGACGTA GTACCGCTTA TAGACCGACT AGTCCACCAT GACCGAGTAC AAGCCACAG
 GGCCTTGCC ACGTACCTT GCGCTTAAGG GGCACGGTTTC TCACTGTGATT CTCGGGGAT ATCCTGGCTGA TCAAGTGGGAT CTCGGCTCATG
 501 TGCCTCTCGC CACCCGCGAC GACGTCGAC GACGTCGAC GACGTCGAC GACGTCGAC GACGTCGAC GACGTCGAC GACGTCGAC
 ACGGGAGGC GTGGGGCTCG CTGCAAGGGGG CCGGGCATGC GTGGGGCATGC
 601 CCACATCGAG CGGGTCAAGC AGCTGCAAGA ACTCTCTCTC ACCTGCGTGC GGCTCGACAT CGGCAAGGTG TGGGTGGGG AGCACGGGC
 GGCTAGCTC GCGGAGTGGC TCGACCTCT TCGAAGGG TCGGGCGAAC CCGAGCTGTA GCGTGTCCAC ACCAGGCC TGTGCGCG
 701 GTCTGGACCA CGGGGAGRG GTCTGGAGC GGGGGCTGT TCGCGGAGAT CGGGCGGC ATGGGGAT TGAAGGGTTC CGGGGTGGCC
 CAGGCTGGT GCGGCTCTC GCACTCTCGC CACGGCTCTA CGGGGCTCTA CGGGGCGACA CGGGGCTCTA CGGGGCGACA CGGGGCGACA
 801 AGATGGAAGG CCTCTGGGG CGGGCACGGC CCAAGGGAC CTCGGCCACCG TCGCTGGTTC CGGGCAC
 TCGACCTCTCC GGTCGGCGC GGGCACCGAG GGGCACCGAG AGGGGTGEC AGGGGAGG CGGGGTGGT GTCGGCG
 901 CGCCGCTCGT CTCCCCGGAG TGGAGGGCGC CGAGGGCGC GGGGTGGCG CTTCCCTGGA GACCTCCACC TCCCTCTCTA CGAGGGCGC
 GCGCCAGCAG GAGGGCCCTC ACCCTCGCCG GCTCGGGCG CCCACCGGC GGAAGACCT CTGGGGTGG AGGGGAGAT GTCTGGCGAG
 1001 GGCTTCACCG TCAACGGCGA CTTGGAGTGC CGCAAGGRCG GGGCGACCTG GTGCTGGGT TAACGCTCC CTCCTCTAAAG
 CCCAACTGGC AGGGGGGGT GCACTCAGC GGCTCTCTGG CGCGCTGGAC CACGTACTGG CGCTTGGGC
 1101 CTATGCTATT TTAATGACCC ATGGGACTTT TGGCTGGCTT AGATCCCTT GGCTTGGTTA GAAAGCAGT ACATTAATA CATAACCTTA
 GATACTTAA AATATTCTGG TACCTGAA ACGACGAA TCTAGGGAA CGGARGAAAT CTGGGTGCA TGTAAATTAT GTATGGAAAT
 ACATACGATT TATGTGACAC TATAGATAAC ATCCACCTTG CCTTCTCTC CACAGGTGTC CACTCCCGG TCCAATGCA CTCCTCTT
 TGTATGCTAA ATCCACTGAA ATATCTATT TAGGTGAAAC GGAAAGAGAG GTGGGGTGC AGGTGACGT GGAGGCAAGA TACCTAA
 1201 TCCCCGGGG ATCCCTCTAGA GTCAACCTGC AGAACCTTG ATGGGGCCA TGGCCCAACT TGTATTGTC AUCCTATAAA GGTGTTATG
 AAGGGGGCCC TAGGAGATCT CAGCTGGAGC TCTTCGAAGC TACCGGGGT ACCTGGTTGA ACAATAAAGC TUGAATAATAA CCAATGTTAA
 1301 TCCCCGGGG ATCCCTCTAGA GTCAACCTGC AGAACCTTG ATGGGGCCA TGGCCCAACT TGTATTGTC AUCCTATAAA GGTGTTATG
 AAGGGGGCCC TAGGAGATCT CAGCTGGAGC TCTTCGAAGC TACCGGGGT ACCTGGTTGA ACAATAAAGC TUGAATAATAA CCAATGTTAA

Figure 2-1

1401	CATCACAAAT TTCAACACATAA AANGCATTTT TTCACTGCAAT TCTAATTTGTG 5' GTTCTGTCAA ACTCATCAT GTATCTTATC AT'G'CTGGAT' CUAATCGG'GA	AAATAGCCG CGTCGTGTTA AAGTGTATAT TTGTTAAACAA AGTGTGACGTA AGTCAACAC CAAACAGTTT TGTGAGTTA CTAAGAATAG TACAGACCTA CCTAGCCCTT'
1501	TTAATTGGCG GCAGCACCAT 'GCCTGTAAT AGACCTCTGAA AGAGGAACTT GTTCTGAGGC CCAATCTGAA CAACTGAGCTT TCTCCTGAA CAAATCTGAG GAGACTCCG CCTTGTCTGG 'TCGACACTT AGACACATC	AAATAGCCG CGTCGTGTTA AAGTGTATAT TTGTTAAACAA AGTGTGACGTA AGTCAACAC CAAACAGTTT TGTGAGTTA CTAAGAATAG TACAGACCTA CCTAGCCCTT'
1601	TTAGGGTGTG GAAAGCTCCC AGGCTCCCCA GCAAGCCAGAA GTATGCCAARG CATGCACTTC AATTAGTCG CAACCAAGGTG TGGAAAGTCC CCAGGTCTCC	AAATCCACAC CTTTCAGGG CGCAGGGGT CATAGCTTC GTCAGTGTAG TAAATCAG CATCCGCC C'AACTCCGC CCAGT'TCCTC
1701	CAGGAGCAG AAGTATGCAA AGCATGCACTC TCAATTAGTC AGCAGCCATA GTCCCGCCC TAACTCCGC CATTCCGCC C'AACTCCGC CCAGT'TCCTC	GTCTGTCGTC TTCACTACGTT TCGTGTGTTA CAGGGGGG ATTGAGGGG GAT'YAGGGG GGT'CAAGGGG
1801	CCATCTCCG CCCCATGGT GACTTATTTT TTATTTATAT GCAGAGGGCG AGGCCGCCTC GGCTCTGAG CATTCCAGA AGTATGTAGGG AGGCT'TCCTT'	GGTAAAGGGC GGGPACCGA CTGATTAAA AAATAATAA CGTCTCCGC TCCGGGGAG CGGGAGCTC GATRAAGGTCT TCACTACTCC TCGAAAAGAA
1901	GGAGGCTTAG GCTTGTGCAA AAAGCTGTTA CCGTGGCGG CGCCTTAATT AAGGGCGGCC ATTTRAATCC TGCAGGTAAC AGCTTGGCAC TGGCCCGT'G'	CTCTGGGATC CGAAACGTT TTTCGACAT GGAAGCTGCC GGGGAAATTAA TTCCGGGGG TAAATTAGG AGCTICATTG TCGAACCGTG ACCGGCAGCA
2001	TTTACACGT CGTGTACTGG AAAACCCCTGG CGTTACCCAA CTTATCAGC TTGAGCACAA TCCCCCCTTC GCAAGCTGGC GATAATGGA AGCAGGACCC CGTGTGGGT' GAATTAAGGGC GCACTGACCC TTGTTGGGAC	AAATGCTGCA GCACTGACCC TTGTTGGGAC GCAATTTGT AGGGGGAAAG CCGTGTGGCT CTTATGCGT CTTATGCGA CGTGTGGCTT TCT'CCGGGGU
2101	ACCGATGCC CTTCCCAACA CTGCGCTGAC CTGCGTGGCG AGTGGCGCTT GATGGGTAT TTCTCTTAA CGCATCTTG CGCATCTTG CGGAAATTCA CACCCGATAC	TGGCTAGGG GAAGGGTGT CAACGGATCG GACTTACCGC TTACCGCATA AAAGGAAAT GCTGAGACAC GCCATAAAAGT' GTGGCGTATG
2201	GTCAGACAA CCATACATCG CGCCCTGTAG CGCGGCATTA AGCGCGGGGG GTGTGGGTG TACGGGAGC GTGACCGCTA CACTGCCAG CUCCTCTAG	CAGTTGCGTT GTGACATGC GCGGGACATC GCGCGGTAT TCGCGTGGCG CACACCA CAATGGCTG CACTGGCGAT GIGACGCTC GCGGGAATGCG
2301	CCCGCCPT TCGCTTCTT CCCTCCCTT CTGGCCACGT 'TCGGCCACGT ATTTGGTCA ACCAGTGTCA AGTGGCCCAT CGCCCTCTTA GACGGTTTT CGCCCTTTGA CGGT'GGGTC	GGGGGAGGAA AGGCAAGAA GGGAAAGGAA GGGGGCGAA AGGGGGCAAGT CGAGATTAG CTGCGTAAAGA GGGGGGAAACT GCAACCTCTAG
2401	CTTACGGCA CCTGCCACCC AAAAACCTG ATTTGGTCA ACCAGTGTCA AGTGGCCCAT CGCCCTCTTA GACGGTTTT CGCCCTTTGA CGGT'GGGTC	GAATGCGGT GGAGCTGGG TTTCGAAAC TAACCCACT ACCAGTGTCA AGGCGGGTCA AGGGGGCAAGT CGAGATTAG CTGCGTAAAGA GGGGGGAAACT GCAACCTCTAG
2501	CACTTCTT TAACTGGAC TCTCTGTCAA MACTGCAACA ACACTCAAC CTATCTGGG CTATTCTTT GATTATPAAG GGAT'YTGCC GAT'YCCGGC	GTGCGAGAAA TTATCACCTG AGAACAGGT TTGACCTGTG GATAAGCCC GATAAGAAA CTAATAATTC CCTAARACGG CTAAGGCCG
2601	TATGGTAA AAAATGAGCT GATTACAA AAATTTACG CAATTPTAA CAAATATAA AGCTTACAA TTATATGTG CTTTATGTG CACTCTCGT ACAATCTU'	ATACACATT TTTCATCGA CTTATGTT GTTATATG CTTTATGTG CTTTATGTG CTTTATGTG CTTTATGTG CTTTATGTG CTTTATGTG CTTTATGTG CTTTATGTG
2701	CTGATGCCGC AGATGTAAGC CAACCTCGCT ATCGCTAGT GACTGGGTA TGGCTGCGCC CGAACACCGG CAAACACCGG CTAAGCGCC	GACTACGGGCT TTACAGACAA CCTGTCAGCA TAGCGATGCA CTGACCCACT ACCGACGGG GGCTGNGGC GGTGTTGGGC GACTGGGG
2801	TGTCATGCC CGGCTTCGCC TTACAGACAA CCTGTCAGCA TAGCGATGCA CTGACCCACT ACCGACGGG GGCTGNGGC GGTGTTGGGC GACTGGGG	ACAGACAGGG GCCCTAGGGC AAATCTGTT CGACACTGGC AGAGGCCCTC GACCTACACA GTCTCCAAA GTGGCGTGTAG TGGCTTTSGG CGCTCGTCA
2901	ATTCCTGAG AGCAAGGGC CTGCGATAAC GCTATTTT ATAGGTTAA GTATGATAA TAATGTTTC TTAGAGCTCA CGT'GGCTAC' 'TTTCTTAA	TAGAACCTTC TGCTTCCCG GAGGACTATG CGTAACTAAAGA ATTCACAAAT CAGTACTTAACTTAAAGGAAATGAA ATTCACAAATGAA
3001	TGTCGGGGA ACCCTTATTG TTATTTTCTTAA CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC	TGTCGGGGA ACCCTTATTG TTATTTTCTTAA CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC CCGTGTAAAC

Figure 2-2

Figure 2-3

4801 GTTATCCCCCTT GATACCGTAT TACCCGCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACAGACCGAGC GCACCCGAGC ACTGACCCGAGC
CATAGGGGA CTAAGACACC TATTGGATA ATGGGAAA CTCACTCGAC TATGGGGAC GGGGTGGCTT TGCTGGCTCG CGTCGGCTAG TCACTGGCTC
4901 GAAAGGGAAAG AGCGCCCAAT AGCCTAACCG CCTCTCCCCG CGCGTGGCC GATTCAATTAA TCCAGCTGGC ACGACAGGTG TCCCGACTG
CTTCGGCTTC TCGCGGGTTA TGGGTTGGC GGAGAGGGGC GCGCRAACGGG CTAAGTAAATT AGGTGACCG TGCTGTCCAA AGGGCTGACCT TTTCGCCCCGT
5001 GTGAGGGCAA CGCAATTAAAT GTGAGTTACC TCACTCATTA GGCACTCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTTGGAA TTGTTGAA
CACTCGCGGT GCGTTAATTA CACTCATGG AGTGAAT CGTGAAT CGTGGGGTC CGAAATGTGA AATAGAAGG CGAGGATAC AACACACCT AACACTCGCC
TATTGGTTAAA GTGTGTCCCTT TGTCGATACT GGTTACTAATG CTTAATT

>length: 5147

Figure 2-4

Figure 3
psV.ID
length: 5171 (circular)

1 TTCGACCTCG CCCACATG ATTATTGACT AGAGCTGATC GACAGCTGTG GAAATGGGT CAGTTAGGT GTGGAAAGTC CCAGGCTCC CCAAGAGGA
 AGCTCGAGC GGGCTGAACT TAATAGCTA TCTCGCTAG CTGTCAGAC CCTACACCA GTCATCCCA CACCTTCAG GGTCGGAGG UGTCGTCGGT
 101 GAGATGAA AAGATGCACT CTCATTAGT CAGCACCAG GTGTGGAAAG TCCCGAGG CCCCCAGGG CAGAGTATG CAAGAACCC ATTCCTGATC
 CTCATAGT TCTGTAAGT GAGTTAAGT GTCCTGTTTC AGACCTTTC GACACCTTC AGGGTCCGA GGGGTGTC GTCATAC GTCATAC
 201 GTCAGCAACC ATAGTCCCGC CCTRACTCC GCCCCATTC CGCCCTAATC CGCCCATTC CGCCCATG GCTGACTAT "T"TTT"TTT"AT
 CAGTCGTTG TATCAGGGGG GGGTGTAGG CGGGTGGAC GGGGATTAAGA GGCGGGTAC CGACTGATTA AGAAAATAA
 301 TATGCAAGGG CGGATGGAA CGGGGATTC CCGGCCAAG AGTGGATGAA GTACGGCTA TAGAGTCTAT AGGGCACCC CTTGGCTTA GAGAGAATA
 ATAGTCCTCC GGCTCGGGG GAGCTTAAGG GGCTCTGGC ACCTAACCTT GGCTCTGGC
 401 CCGGGRAGGG TGGATGGAA CGGGGATTC CCGGCCAAG AGTGGATGAA GTACGGCTA TAGAGTCTAT AGGGCACCC CTTGGCTTA GAGAGAATA
 GGCCCTGGC ACCTAACCTT GGCTCTGGC
 501 AGCCATGAT TTATCCCGS GTGGCCTATC GTGGTGCCTA TTGAGCTGA TGTGCGGGT GTGCCAATAT ATGGGATG AGACATTAATC
 TCGGATCTTA AAATAGGGGC CTCGGTAGTA CCAAGCTGGT AACCTGATG AGGAGGGCA CAGGGTTTAAC CGTCTCTGCGC
 "DHR ATG
 601 TGCCCTGGC TCGGRACSC GTCAGATGAC TTCCBAAAGAA TGACCCAAAC CTCTTCATG GAAGGTAAAC AGAAATCTGGT GATTTATGCT
 AGGGAGGGC AGTCCTGGC CAACTTCATG AAGGTTCATG ACTGGTGTG CAAATACCCAA TCTCTGAG
 701 GGTTCCTCCAT TCCGTGAGG ATAGGACCTT TAAGGAGAG AATTATATAA GTTCTAGTA GAGAACCTTA AGAACACCA CGAGGAGTC ATTTTCTGGC
 CCAAGAGGTA AGGACTCTTC TTAGCTGGAA ATTTCCTGTC TTATTTAT CAGAGCTAT CTCTTGAGTT TCTTGTGGT GCTCTCGAG TAAAGAGAC
 801 CAAAGTTG GATGATGCC TTAGCTATT TGAAACCG GAATGGCAA GTAAAGTAGA CATGGTTGG ATAGTGGAG CAAATACCCAA TCTAGCCTC
 GTTTCBAC CTCATAGCA ATTCTGAA ACTTCCTGGC TTARACCTT CTTACCTT GTRCCBAACT GTCAGCTC CTCAGAGACA ATTCGTCCTC
 901 GCCATGATC AACCAGGGCA CCTTAGACT TTGTGCAAA GGATCATGCA GGATTTGAA ATGTGACAGT TTTCCCGAA ATTCGTTTG GGGAAAYATA
 CGGTACTTAG TTGGTCCGGT GGARTCTGAG AAACATGTT CCTAGTACTT CCTTAACCT TCACGTGTC AAAGGGTCT TTAACTRAAC UCTT"TTT"AT
 1001 ABCCTCCTCC AGATAACCA GGCGTCCCT CTGAGGTCCA GGAGGAAAA GGCATCAAGT ATAAGTTGA ATGTCAGGAG AACAAAGACT AACAAAGAA
 TTGGAGGGG TTCTATGGGT CGCGAGGAA GACTCAGGT CCTCCCTTT CGGTACTCA TATTCAACT TCAGTGTCTGTA TCTTCTGTA "T"TCCTCTG
 1101 TGCTTCAG TTCTCTGCTC CCCCTCTAAA GCTATGCTT TTATATAGAC CATGGGACTT TGTGTTGGCTT TAGACCCCTT TGGCTTGGCTT
 ACGAAAGTC AAGAGACGAG GGGGAGATT CGATCGTA AAATATCTG GTACCGTA AACGACCCAA ATCTGGGGA ACCGAAGCAA TCTTGGCC
 1201 TACATATAT AGATAACCTT ATGTTATATAA CACATAGATT TAGGTGACAC TATAGATAA CATCCACTT GCCTTCTCT CTCAGTGT GATTTCTA
 ATGTTATATAA TGTATTGAA TACTTAGT GTGTTCTTA ATCCACTGTT ATTCAGTAA GTCAGTGT GATTTCTCTA GTCAGTGT
 1301 CAACTGACCC TCGTTCTAT CGATTGAAATT CGCTCTAGAT CCCGGGGGA" CCTCTAGAT CGACCTGAG AAGCTGCGC GCAATGGCCC
 GTTGAACCTGG AGCCAAGATA GCTAACTPAA GGGGSCCTTA GGAGATCTCA GCTGAGACGTC TMCGAACGG CGG"ACCGG TMCGAACGG
 1401 TAATGGTTAC AAAATAAGCA ATAGCATCAC ATATTCAAA ATAAGGAT TTTCCTACT GCTATTCTAGT TGTGTTTGTG TGTGTTTGTG
 CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT CAAATCTCAT

Figure 3-1

ATTACCAATG TTATTCTGT TATCCTAGTG TTAAAGTGT TTATTCTGT AAAAAGTGA CTTAGATCA ACCAAARCA GTT"TGAGTA GT"TGATAGA
 1501 TATCATGTCT GGATCGATCG GGAATTAACTT CGGCCGAGCA CCATGGCTCG AAATAACCTC TGAAGAGGA ACCTCGTTAG GTACCTCTG AGCGGAAAGC
 ATTGTACAGA CCTAGCTAGC CCTTAATTAA GCGGCGTGT GGTACCGAC TTTATGGAG ACTTCTCCCT TGAACCAATC CATGGAGAC TUCGCCCTTC
 ^sv40 origin

1601 AACACAGCTGT GGATGTGTT TCAGTTAGG TGTTGAAAGT CCCAGGCTC CCAGCAGGC AGAAGTATGC AAAGCATGCA "TCACCAATTAG TCTCAATTAG
 TTGGTGGACA CCTTACAGAC AGCTAATCCC ACACCTTCA GGGGTCGAG GGGTGTCCG TCTTCATAGC TTCTGACGT AGAGTTAATC AG"CGT"TCGTT
 1701 GGTGTTGAA GTTCCAGGAGC GCGAGAGT GCAGAGCTG CTCAGGAGT CTCAGGAGT CTCAGGAGT CTCAGGAGT CTCAGGAGT CTCAGGAGT
 CCACACCTT CAGGGCTCGC AGGGTGTGAG GGGGGTCA AGGGGGTAA CGCAGCTGTT AAAAALAAATA AATACGCTTC CGGCTCCGGC GAGCUCGGAG ACTCGATAA
 1801 CCCCTTAACT CGGCCCAAGT CCCGCCATTC TCCGCCATTC GGCTGAGTAA TTTTTTAT TTATGAGAG GCGAGGGCC CMCGGCTC TGACCTATTC
 CGGGGATTGA GGGGGTCA AGGGGGTAA CGCAGCTGTT AAAAALAAATA AATACGCTTC CGGCTCCGGC GAGCUCGGAG ACTCGATAA
 1901 CAGAAAGTGT GGGAGGGCTT TTTGGAGGC CTAAGCTTT GCAGGTTTGCAG GTTACCTCGA CGGGCCCTT AAT"AGGCG CGCCATTAA ATCCCTGAGC
 GTCTTCATCA CTCTCCGA AAAACCTCG GATCCGAAGA CGTPTTGAG CAACTGGAGCT CGCCGGGAA T"AAATTCCCC GCGGTAATT TAGGAC"TC
 ^start pUC18

2001 TAAACAGCTTC GCACTGGCCG TCGTTTACA AGTCGTGAC TGGGAAACC CTCGGCTTAC CCAACTTAAT CCCTGGCGA CAACTCCCC
 ATTGTGAAC CGTGACGGC AGCAGAAATGT TGAGCAGTC ACCCTTTGG GACCCCAATG GGTTGAAATG GCGGAACCTC GTCAGGGGG
 2101 TGGCTTAATA GCGAAGGGC CGGCACCGAT CGCCCTTCCC BACAGTTGCG TAGCTGAAAT GGCAGATGGC GGCTGATCG GTCATTTCTC
 ACCGCTATTAT CGCTTCTCCG GGGTGGCTA GGGGAAGG TGTCAACGC ATCGCTACTA CGCGCTTACCG CGGACTAGGC CATAAAAGAG
 2201 TGTGGGTAT TTACACCCG ATACGTCAA GAAACAGTAG TACGGCCTT GTAGGGCCG ATTAAGGGC GCGGGTGTGG TGGT"AGCG
 ACACCCATA AGTGTTGGC TAGTGCTTT CGTGGTATC ATGCGGGGA CATCGCGG TAAATGGC CGCCGAGCC ACCAATGCC
 2301 GCTAACATG CCAGGGCCCT AGCGCCCGCT CTTTCGCT TCTTCCTC CTTCTCGCC ACGTTGGCG GCTTCCCCG "CAACCCCTA AAT"CGG
 CGATGTGAC CGTCGGGAA TCGCGGGAA AGAAGGGAG GAAAGGGG GAAAGGGG AGTTGGAG ATTCGAGAT TTTGGCCCG
 2401 TCCCTTGTAG GTCCGATT AGTGCTTAC GGACCTCGA CCCCCAAA CTGATTTGG GTGATGGTC AGTGTGGG CCATCGCC" GATAGAGC
 AGGAAATCC CAAGGCTAA TACGAAATG CGTGGAGCT GGGTTTTT GAACTAACC CACRACAG TGCATCACCC GTTGGGGAA
 2501 TTTCGCCCCT TTAGCTGG AGTCCAGTT CTTAAATAGT GGACTCTGT TCCAAACTSG AACACACTC AACCTCTATCT CGGGCTATTC
 AATGGGGAA AACTCACCC TCAAGTCAA GAAATTATCA CTCGAGACAA AGTTTGACCT TTGGTGTGAG TTGGGATAGA GCGCGATAAG AAACCTAAG
 2601 TAACGGATT TGGCGATTTC GGCTPATGG TTAAAGTGT AGCTGATTA AGRAAAATT AAGCCGATT TTAACAAAT ATTACGTTT AGAAATT"TA
 ATTCCCTAA AGGGCTAAAG CGGATAACCA ATTCTTAC TCGCTTAAT TGTCTTAA TTGCGCTTA AATTTTTTA TAATGCTAA TGTCTAAAT
 2701 GGTCCTCTT CAGTACAATC TGCTCTGATG CCGCATAGT AAGGCCACTC CGCTATCCCT ACOTGACCTGG GTCTGGCTG CGCCGCCAACAA
 CCAGTGTGAG GTCTGTTAG AGAGACTAC GGCGPATCRA TTGGGTGAG GCGATAGCA TGCACGTGAC CAGTACCGRC GCGGGCTGT
 2801 CCCGCTGAGC CGCCCTGAGC GGCTGTCTG CTCGGGGCAT CGCCTTACAG ACAGGCTGT ACCGTCCTCG GGAGCTGCT
 GGGGACTGTC CGGGGACTGTC CGAGACAGAC TGGCAAGTC TGGCAAGTC TGGCAAGTC CCGCAAGTC CAGTACGTG
 2901 CTCACCAA ACCGGGAGG CAGTATTCTT GAGAGCRAA GGGCCCTCGT ATACGGCTAT TTTTATAGT TATGTCATG ATTAATATCG
 GTRGTGGCTT TGGGGCTEC GTCTAAGAAT CTCTGCTT CGGGGGATA AAAATATCA ATTACGTC TATTAATTCAGC TATTAATTCAGC
 3001 GTCAAGGTGCG ACTTTGGG GAATGTGCG CAGTACAGAC TGGCAAGTC GGGCAAGTC GGGCAAGTC
 CAGTCCACCG TGAAGGGCC CTTACACCG
 3101 TAAATGCTC AATATATTG AATAGGAG AGTATGAGA TTCAACATT CGGTGTCGCC CCTTATCCCT TTTTGGGC AT"TTTGUCC" CTCGTTT
 ATTACGAG TTTTCCCTC TCATACTCA AAGTGTAA GGCACRGCGG GAAATGGGA AAAAACCGCG TAAACGGGA GAGCAAAAC

Figure 3-2

3201 CTCACCCAGA AAGCTGGTG ATGCTGAGA TCAAGTGGG GCACGAGTGG GTTACATCGA ACTGGATTC AGATCCCTTGA AGATGGGTA AGATCCCTTGA
GAGTGGTCT TTGCGACCA TTTCATTTTC TAGACTCTTC AGTCATCCCA CTCGTCACCA CAATGATGCT TGACCTAGAG TTGTCGCCAT TCTAGAACCT
3301 GAGTTTCGC CCGAGAAC GTTAAAGT GATGAGGACT TTGAAAGTTC TGCTPATGTT CGCGTATTAA TCCCGTGTAG AGCGGGGCA AGCGAACCT
CTCAAAAGCG GG GCFCTTG CAAAGGTT CTAAGCTGAA AATTCAAG AGGTACACC GCGCCATAAT AGGGCACTAC TGGGCCGCT TCGUGTTGAG
3401 GGTGCGCA TACACTATT TCAGAATGAC TGGTTGAGT ACCTACCAAGT CACAGAAAG CAUCCTACGG ATGGCATGAC AGTAAGAGA TTAAGCAGT
CCAGGGCGT ATGTTAAAG AGTCTTACTG ACCAAGTCA TGAGTGGCA GTGTCTTTC TGTAGATGCC TACCGTACTG TCAATTCTT AATAGCTCAC
3501 CTGGCTAAC CAGAGATGTAACAGT AACTCTGG CAACTTACTC TGTGAGGAC ATGGAGGAC CGAAGGAGT AACCGCTTTT T'GSCACAA 'TGGSCATTA
GRCGSTATTC GTAACCTACTA TTGTGAGGCC GGTGAAAGA AGACTGTTC TAGCTCTG GCCTCTCGA TTGGCGAAA AACGTGTGT ACCCGCTAGT
3601 TGTAACTCGC CTGTATGTT GGGRACGGGA GCTGAATGAA AGCCTACCAA AGCAGAGGCG TGACACCAAG ATGCCAGAG CAATGGCAAC AACGGTGC
ACATGAGCG GAACTGCAA CCCTTGCGCT CGACTTACTT CGGTATGGTT TGCTGCTCGC ACTGGTGTGC TAGGGTGTGC GTTACCGTGTG TTGCAACGCG
3701 RAATTTAA CTGGGCACT ACTTACTCTA GCTTCCCGC ARCAATTAAAT AGACTGATG GAGGGGATA AAGTGGTGAAGG ACCACTTCTG CGCTCGG:CC
TTTGTAAATT GRCGCTTGA TGAATGAGT CGRAGGGCCG TTGTTAAATA TCTGACATAC CTCCCGCTAT TTCAAGCTAC CTACCATCTG GCGAGCCGGG
3801 TTCCGGCTGG CTGGTTTATT GCTGATAAT CTGGGCGG TGAGGGCGG CTGGGCGG TCTCGGCTA TCATGGCAGC ACTGGGCCA GTGGTAAGC CC'TCCGGTAT
AAGGCCGACC GACCAAAAT GCACTATTA GACCTTCAAGC ACTTCGCGC AGAGGCCAT AGTAACTGCG TGACCCCGGT CTACCATCTG GGAGGGCATTA
3901 CGTAGTTATC TACAGGAGGG GGAGTGGGG AACTATGGAT GRACGAAPATA GACAGATGCC TGAGTAGGT GCCCACTGTA T'YAAGCATTT GTAACCTGCA
GCHMCAATAG ATGRCGCTGCC CTCAGTGGC TTGATACCTA CTTCGTTAT CTGCTAGCG ACTCTATCCA CGGAGTACTA AATTCGTAAC CATTGACAGT
4001 GRCGAGTTT ACTCATTTAT ACTTGTAGT GATTTAAAC TTCAATTAAAGT ATTAAAGG ATCTAGGTGA AGATCCCTTT TGATAATCTC ATGACCAAAA
CTGGTTCAAAGA TGAATTTATA TGAATCTAA CTAAATTGAG ATGTAATTTAC TAAATTCTC TAGTCCACT TCTAGGAAA ACTATPAGAG TACTGGTTT
4101 TCCCTTAACG TGAGTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAGAG ATCAAGGAT CTCTGTGAA TCCCTTTTTT CTGCGCTAA 'TCTCTGTT
AGGRRATTCG ACTTAAAGC AGGTGACTC GCAGTCTTTC TACTTCCCTA GAAGACTCT AGGAAAGAAA GRCGCGATI AGAGGACGAA
4201 GCAARACAAA AAACCCACCGC TACCAAGCGG GGTITTTG CCGGATCAG AGCTTACACAC TCTTTTCCG AAGGTAATCTG GCTTCAGCAG AGGGCAGAYA
CGTTGTTT TTGTTGCGC ATGGTGCCTA CCAAACARAC GGCCTAGTC TGATGGTGT AGAAAAGGC TTCAATTGAC GGAAAGTGTCTC TCGGCTCTAT
4301 CCAATACTG TCTTCTAGT GTAGCGTGTAG TTAGGGCAC ACTTCAGAA CTCGTAGAA CCCTCTACAT ACCCTGCTCT GCTTAATCCTG T'ACCAGTGTG
GTTTTRTCG AGGRRATCA CTCAGGATCA CTCAGGATCA ATTCGGTGTG TGAAGTCTG GAGACATGT GGCGATGT TGAGGCGAGA CGTATTAGAC ATGGTCACCC
4401 CTCGCTGCCAG TGGCGATAAG TTGTTGCTTA CGGGTTGGA CTCAAGAGCA TAGTTACGG ATANGCCGA GCGGTCGGGC TGAAACGGGG GTTCTGTC
GAGGACGGTC ACCCTCTATTG AGCAGCAAAAT GGCCCTACCT GAGTTCTGCT ATCAATGCTC TTTCGCGGT CGCAAGGCCG ACTTSCCCC CAGGACGCTG
4501 ACAGGCCAGC TTGGAGGCCA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGGATTG AGANAGGCCG AGCTTCCCG AGGGGAAAG GCGGGAGAG
TGTGGGTCG AATCTCTGTT GTGAGATGTC GTTGTGCT ATGGATGTC CACTCTAAC GCTTGTGCG TGCGAAGGGC TTCCCTCTTT CGGCCCTG
4601 TATCGGGTAA GGCGCAGGGT CGGACACAGGA GAGGCACCGA GGGAGCTTC AGGGGGAAAC GCCTGGTATC TTATATAGTCC TGTGGGTTT CCGGCACTC
ATAGGCCATT CGCGTCCCA GGCTGTGCT CCCTCGAAGG TCCCTTTTG CGGACCATAG AAATATCAGG ACAGCCAAA GCGGTTGAGA
4701 GACTTGAGCG TCGATTGTTG TGATGCTGTT CAGGGGGGG GACCCATAGG AAAAACGCCA GCAACGCCG CTTTTTACGG T'CTCUTGGCCT T'TTGCTG
CTGAACTCTC AGCTAAAAAAC ACTAACGAGCA GTTCCGGATCC CTCGGATCC TTTTGGGT CTTTGCGCG GAAAATGCC AAGGACGGA AAACGACGG
4801 TTTTGCTCAC ATGTTCTTTC CTGGTTTACCT CCGTGTATTG CCGTGTATTG GTGGATAACC GTATTAACCG CTTTGAGTACGG CTCCTCGAG CTCGACG
AAAACGACTG TAAAGAGAAG GAGGCAATAG GGGGACTAAGA CACCTATGG CTCAGTGGCA GAAACTCAGC GAAACTATGG CTCAGTGGCA GTCGTTG
4901 GAGGCCAGCG AGTCAGTGTAG CGAGGAAAGC GAGGAGGCC GAAATAGCAA ACCTAGCAA ACCGCTCTC CCCGGCGTT GCGCGATCA TTAATCAGC
AAACGACGG

Figure 3-3

CTCGGCTCGC TCACTGACTC CCTCCCTTCGC CTTCTCGCG GTTATGGCTT TGGGGAGAG GGGGGCAA CGGGCTAAGT AATTAGGTG ACCGGCC'U
5001 GGTTTCCGA CTGGAAAGCG GGCAGTGGC GCAACGCAAT TAATGTGAGT TACCTCACTC ATTAGGACCC CGAGGCTTA CACTTATUC "UCCGGCTC
CCAARGGCT GACCTTCGC CGGTCACTCG CGTTGGCTTA ATTACACTCA ATGGAGTGAG TAATCCGGG GGTCCGAAT GTGAAATAACG AAGGCCGAG
5101 TATGTGTG GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT ATGCCATGA TTACGAATTA A
ATACAAACACA CCTPACACT CGCPACACT CGCCTATGT TAAGTGTG AACTGTTGCA TACTGGTACT AATGGTTAAT T
>length: 5171

Figure 3-4

Figure 4
PSV, IPD
length: 5712 (circular)

Figure 4-1

Figure 4-2

Figure 4-3

CTATTCGGCG TGCCAGGCC GACTTGGCCC CCAAGCACCT GTGTGGGCT GAACCCGCT TGCATGGATGT GGCTTGACTC TATCGATGTC GCACTCGTA
 5101 GAGGAAAGGC CAGCCTTCC GAGGGAGAA AGGGGACAG GTATCGGTA AGGGGAGGG TCGGAGACGG AGGGCGACG AGGGAGCTC CACGGGGAAA
 CTCCTTCGGC GTGGCGAAGGG CTCGGCTCTT TCCGGCTGTC CTCGGCTCC AGCGGCTCAT TCGGGCTCC AGCGCTGTC TCCCTCGAAG GTCCCCCTT
 5201 CGCTGTGAT CTTATAGTC CTGTCGGGT TCGGCCACCTC TGACTTGACG CTCGATTTT GTGATGCTG TCGGATTTT GTGATGCTG TCGGGGGGC GGAGCCATG
 GCGGACCATTA GAAATATCG GACAGCCCA AGCGGTGGAG ACTGACTCG CAGCTAAAAA CACTACGAGC AGTCCCGCG CTCGGATAAC
 5301 AGCAGCGGG CCTTTTACG GTTCCTGGCC TTTCGCTCA CTTTGCTCA CAGTGCTCTT CCTGGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG
 TCGTTGGCC GGRAGATGCG CAGGGACCGG AAACGRCGG GAAACGAGGT GTACAGGAAAGA GGACGCRATA GGGGACTAAG ACACCTATAG GCATAATGGC
 5401 CCTTTGAGTG AGCTGATAACC GCGGAAACGAC CGAGGCCGAG GAGTCAGTGA GCGAGGGAGC GCGAGGGAGC CGAAATACGCA AACCGGOCCTC
 GGAAACTCAC TCGACTATGG CGGCGGGGT CGGCGGGGT GCTCGCTG CTCAGTCACT CGCTCCPTCG CGCTCCPTCG CGTATGCGT GTGGCGGAGA
 5501 CCCGGCCGCT TGGCCGATT ATTAAATCCRG CTGGCACGAC AGGTTCGG AGTGGAAAGC GGGCAAGTGAG CGAACGCAA TTACCTCATC
 GGGCGGGCA ACCGGCTAAG TAAATTAGTC GACCGTGTG TCCARAGGGC TGACCTTTCG CCCGTCACTC GCSTTGCGGT ARTAACACTC ARTGGAGTGA
 5601 CATATGCCAC CCAGGGCTT ACATTTATG CTTCCGGCTC GTATGGTGTG TGGAAATGTG AGCGGAAAC AGTGGAAACAGC TATGACCAAC
 GTAATCCGTG GGGTCCGAA 'TGTGAAATAC GAAAGCCGGAG CATAACAACAC ACCTTAACAC TCGCCATATG TTAAAGTGTG TCCCTTGTCG ATACTGGTAC
 5701 ATTACGAAATT AA
 TAATGCTTAA TT

>length: 5712

Figure 4-4

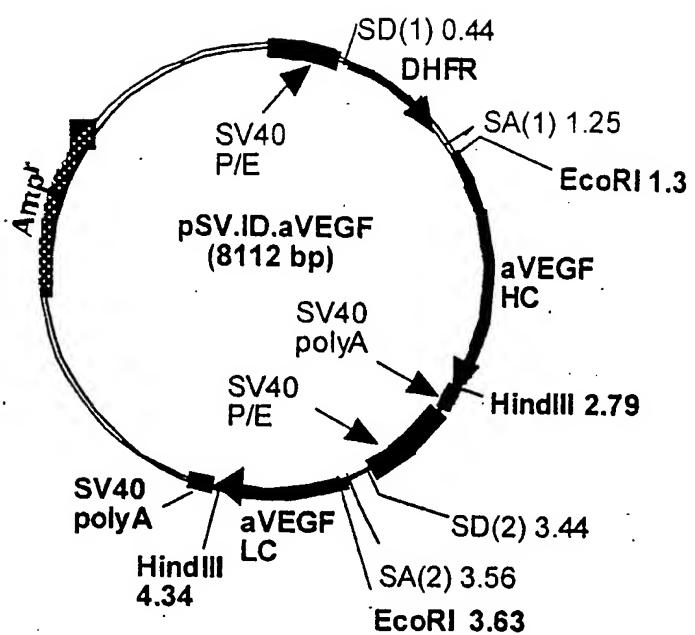


Figure 5. pSV.ID.aVEGF control plasmid

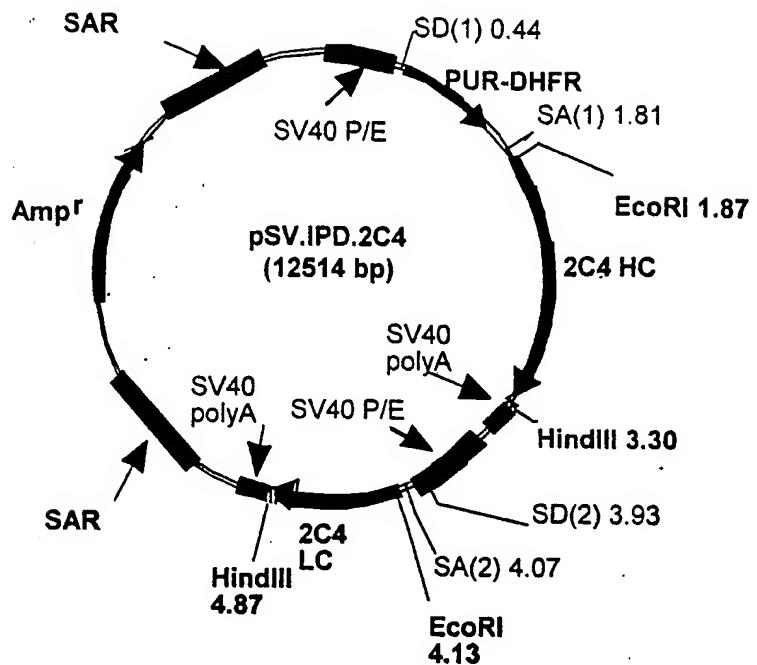


Figure 6. pSV.IPD.2C4

Figure 7
psv.ipd.2c4

length: 12514 (circular)

1 TTGGAGCTG CCCGACATT ATTATTGACT AGACTCGATC GACAGCTGT GATCTGTT CAGTTAGGT GTGGAAGTC CCCAGGTCTC CCAGCAAGCA
 AGCTCGAGC GGGCTGTAAC TATACTGTA TCTGAGCTAG CTCAGCTAC CTTACACACA GTCATCACA CACCTTCAG GGGTCGAGG GTTCGTCGTT
 101 GAACTATGCA AGGATCCT CTCATTAGT CAGCAACAG GTGTTGAAAG TCCCCAGGT CCCACGAGG CAGAAGTATG CAAGCATGCC ATCTCAATTA
 CTCATACGT TTGCTGAGTA GAGTTAATCA GTCGTTGTC CACACCTTC AGGCTTCGA GGGTCCTCC GTCCTCATAC GTTCTGAGG TAGAGTAAAT
 201 GTACGCAACC ATAGTCCGC CCCTAATCTC GCCCATCGCG CCCCTAATCTC CGCCCAAGTC CGCCCATCTC CGCCCATCTC CGCCCATCTC
 CAGTGTTGG TATCAGGGCG GGGATTAGG CGGGATTTAG GGGGATTTAG GGGGATTTAG GGGGATTTAG GGGGATTTAG GGGGATTTAG GGGGATTTAG
 301 TAGGCAAGGG CGAGGCGCTCT GAGCTTATTC AGATGATGTC AGATGATGTC AGATGATGTC AGATGATGTC AGATGATGTC AGATGATGTC
 ATAGCTCTC GGCTCGGGG GAGCCGAGA CTCGATAGG TCTCTCATAC TCCCTCGAAA AACCTCTGGG ATCCGAAAC GTTTCCTCAT CGAATAGGCC
 401 CGGGAAACGG TCCATTGGAA CGGGATTTC CCGTGCGAG AGTCGAGTA GAGCCGCTTA TAGGGCAGCT AGTCGAGCT AGTCGAGCT
 GGCCTTGCC AGTCGACCTT GGCCTAAGG GGCACGCTTC TCACTGATT CAGGGGGAT ATTCGCTGA TCAAGCTGATG TTTCGGTGCCT
 501 TGGCCTCGC CACCGCGAC GACGTCGGCC GGGCCGCTAG CACCCCTGGC GGGCGTTCG CGCGATTAACCC CGCACCTGGC
 ACCGGAGGG GTGGGGCGCTG CTGCGGGGGG CCGGGCATGG CCGGGATGGG GCGGATGGG GCGGATGGG GCGGATGGG GCGGATGGG
 601 CCACATGGAG CGGGTCAACCG AGCTGCAAGA ACTCTCTC ACGGCCCTCC CGCTCGACAT CGGGTGGGG AGCGCCGGC CGCGGTTGGG
 GGTGAGCTC GGCCTAGGG TCAGCTGTTCT TCGAGCTTCTC TCGAGGGAG GGGGGGGGGT TCGGGGAGAT CGGGGGGAT
 701 GTCGCGACCA CGGGGGGGCC CGGGGGGGCC CGGGGGGGCC CGGGGGGGCC CGGGGGGGCC CGGGGGGGCC CGGGGGGGCC
 CAGACCTGGT GCGGCTTCG CCGGCTTCG CCGGCTTCG CCGGCTTCG CCGGCTTCG CCGGCTTCG CCGGCTTCG
 801 AGATGGAGG CTCCTGGGG CGGACCGGCC CGGACCGGCC CGGACCGGCC CGGACCGGCC CGGACCGGCC CGGACCGGCC
 TCTACCTTCC GGAGGACGGC GGAGGACGGC GGAGGACGGC GGAGGACGGC GGAGGACGGC GGAGGACGGC GGAGGACGGC
 901 CGCGTCTGGT CTCCCCGGAG TGAGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC
 CGGCGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC CGGGGGCC
 1001 GGCTTCACCG TOACCGCGAG CGTCGAGTGC CGCAAGGACCC GCGCGACCTG GTGCGTGAAC
 CGGAGTGGC CGGAGTGGC CTCAGCTCG CGGCTCGAC CACGCTACTGG CGGTTCGGG CACGCTACTGG CGGTTCGGG
 1101 TCGTCCCGT GTCCCAAAT TGGGGATG GCAAGAGGAG AGACCTACCC TCCCTTCGAG AGCTCGAGC GTCAGTAC
 AGCAGGGCA CAGGGTTTA TACCCCTAAC CCTTCCTCCG TCGGGATGGG AGGGAGGGC AGTCCTGAG CAACTCATG
 1201 CTCCTAGTG GAGGGAAACG AGAATCTGGT GATTATGGGT AGGAAACCT GGTTCTCCAT TCCCTGAAG AATGACCTT
 GAGAAGTCAC CTCCTCTTGG TCTTGTACCA CTCATACCA TCCCTTCGAG AGGAGGAGTA AGGAGGAGTA
 1301 GTTCTAGTA GAGAAGTCACCA AGAACCCCAA CGAGGAGCTC ATTTCCTGAG TCAAAAGTTG GATGAGCT
 CAAGAGTCAT CTCCTGAGTT TCTTGGTGGT GCTCCCTGAG TAANGAGG GTTTCACAC CTCATACGA
 1401 GTCAGCTGAGA CTCAGGTTGG ATAGTGCGAG GCAGTCTGT TTACCGAGGA GGCATGAATC AGGAGCTCTC
 CATTCTACCT GTACCAAACCT TATCAGCCTC CGTCAGACCA ATGGTCCCTT CGGTACTAG TTGGTCGG
 1501 GGAAATTGAA AGTGACAGT TTTCCAGA AATTGATTTG GGAATAATA AACCTCTCCC AGCATACCA
 GGCTCCCTCT CTCAGGTCACCA CGGAGGAAAN

Figure 7-2

3301 ACCTGCAGAA GCTTCGATGG CGGCCATGG CCAACTTGT TATTCGAGCT TATAATGGT ACAATAAAG CATTAGCATC AGAAAATTCA CAAATAAAGC GTTATTTCG
 TGGAGCTT CGAAGCTTACG GGGGTACG GGTGAACCA ATAACTGTCG ATATTCGAA TGTTTATTC GTATCGTACG TGTTAAAGT
 3401 ATTTCCTCA CTGCACTCTA GTTACATGTT GTCATGTTGAT CTTATGATGAT CTGATGTTG AATTAATGG CTCATGTCG TTTATTCAGC ATGGCTGAA
 TAAATAAGT GRCGTTGATG CAACTACCAA CAGGTTTGG TAGTCTCTA GAATCTCTA GRCCTACCC TTACATTCAGC TACCTGACTT
 3501 ATAACCTGT AAAGAGGAC TTGGTTAGT ACCTTCTGAG CGGGAAAGAA CCAGCTGGG ATATGGTGC ACTTGGGTG TGGAAAGTCC CGAGGCTCC
 TATTCGAGAC TTTCTCTTG ACCAATCTCA TGAAAGACTC CGCCCTTCTCTT GTTCCTCACC TTACACAGC TCATCTCCAC ACCTTTCAGG
 3601 CAGCAGGCAG AGTATGCA AGCATGGATC TCAATTAGTC AGGAAACGG TGTGGBAGT CCCAGGCTC CCCAGGAGGC AGAGATATGC
 GTCGTCCGTC TTCTATGTT TGTCATGTT AGTTAATCG TCGTGTGTC ACACCTTCA GGGTCCGAG GGGTCCGCG TTTCCTACG TTTCCTACG
 3701 TCTCAATTAG TCAGGACCCA TAGTCCCGCC CCTTAACCTGG CCCATCCCGC CCCTAATCTC GCCAAGTCTC CCCATCTC CGCCCATG
 AGAGTAACTC AGTCGTTGGT ATCAGGGGG GGTGAGGCG GGGTGGGGC GGGTCAAGG CGGTAGGG GGGTAAAGG GCGGGTACG
 3801 TTTTATTG ATGCAAGGG CGAGGCCCG CGGGCTTCG AGCTATCCA GAAGTATGAA GGAGGCTTT TTGGAGGACT AGGGCTTTG
 AAAAAAGCTG AAAAAATAAA TACGTCTCCG GCTCCGGCG AGCGGGAGAC TCGTAAAGGT CTTCACTACT CCTCCGAAAA AACTCTCTGA
 3901 CTTATCCGGC GGCCCTGCC CGTGGGAC CGTGGGATCCC CGTGGGAAAGG CGAGGTTCT CAGTCATTC ATGGGGATA TCTCAGTAT
 GATTAAGCCG GCGCTTGCA CGTGGGAC CGTGGGATCCC CGTGGGAAAGG CGAGGTTCT CAGTCATTC ATGGGGATA TCTCAGTAT CGGGTGGG
 4001 TTGAAACGGC GCTACATTA ATACATTAACC TTGGGATCG ATCCTACTGA CACTGACATC CACTTTCT TTTCTCCAC AGGTGTC
 BATCTTGCGC CGATGTTATAT TATGTTATGG AARACCTAGC TAGGATGACT GTGACTGTAG GTGAAAAAAGA AAAAGAGGT
 4101 AACATGACCT CGTTGCGA ATGTTGATCG ATGTTGATCG ACCATGGATG GTGTTGATGT CTTCTCTTTT CTAGTAGCA
 TTGACGTTGA GCACAGGCT TGTATGCAAC CCGACGTTAC TAACTTAAGG TGGTACCTA CCAGPACATA GTAGGAAAAR GATCATCGTT
 4201 AGTACATTCA GATATCCGA 'GACCCAGTC CGCGAGETUC CTTGTCGGCT CTTGTCGGCA TAGGGTCACC ATCACCTGCA
 TCACTGAATG CTATAGCTCT ACTGGTCAG GGGCTCGAGG GACACCGCT ATCCCTGTTAGTGGCTGCT
 4301 ATTGGTGCG CCTGGTATCA AGAARACCA GGAARACTCT GGAARACTCT GATTACTCTG GCTTCCTACG GATACTGG AGTCCCTCT
 TAAACACAGC GGACCTAGT TGTCTTGG CCTTTCGAG GCCTTGATGA CTAATGAGC CGAAGGATGG CTATGTGACG TGAGGGANGA
 4401 GRTCCGCTTC TGGGAGGGT TTCACTCTGA GGTGACCTCT GGTGACCTAT CTTCTGAAAGC GTTGTGTTT ATCTCTTACAC
 CTAGGCAAG ACCCTGCTA RAGTGAAGCT AGAGGTCGGT CTTCTGAAAGC GTTGTGTTT ATCTCTTACAC
 4501 GRTGGACAG GGTACCCGAG TTGGAGATCRA AGGAACCTGT GCTGACCCAT CTGCTCTCAT CTTCCGCCA TCTGATGAGC AGTTGAAATC
 CAAACCTGTC CCATGGTTC ACCTCTAGTT TGCTTGAC CAGTGTGTTA GACAGAAGTA GAAGGCGGT AGACTACTCG TCAACTCTAG ACCTTGACGA
 4601 TCTGTTGTGT GCCTGTGAA TAACTCTAT CCAAGAGGG CCAAGATACA GTGAAAGGTG GATAAGGCC TCAATGGG TAATCTCCAG
 AGAACACACA CGGAGGACTT ATTGAAGATA GGGTCTCTC GGTTCATCT CTTCTGAAAGG AGGTAGGCC ATGGGGTC
 4701 CAGGAGGAGA CAGCAGGAGC AGCACCTACA GCCTGTGAG CACCTGTAG CAGACTACGA GAAACACAA GTCTACGCC
 GTCTGTCTCT GTCTGTCTG TGTGTGATGT CGGAGTGTGCT GACTGTCTC GCTGGACCTC GACTGTCTC
 4801 CCATCAGGGC CTGAGCTCGC CCGTCACAA GAGCTCACAA AGGGAGAGT GTAAAGCTC GATGCCGCC ATGGCCAAAC TTGTTTATG
 GGTAATCCCG GACTGAGCG GCGAGTGTG TCCCTCTCA CATTGAG TACCGGGGG AACRAATAAC GTGGAATATP
 4901 TGGTACAAA TAAGCAATA GCATCACAAG TTTCACATCA TTTACATGCA TTCTPAGTTG GTTGTGCTCA AACTCATCAA
 ACCATGTTT ATTCTGTTT CGTAGTGTAA AAGTGAAGT TTTCGTAAAG AAGTGAAGT TTGAGTGTGTT
 5001 CATTCTGGAA TCGGGGATTA ATTCCGGGCA GCACCATGGC CTTGAAATAG TTAAACCCCT CTGAAAGGG AACCTGGTTA
 GTACGACCT AGCCCTTAAT TAAGCCGGT CGTGGTACCG GACTTATTC AATTTGGGA GACTTCTCC TTGAACTAAT CCATCGTACG
 TCACTGTTCC

Figure 7-3

5101 TCGCCACGCA CARAGTAACT ATTACAATC AGTCACTCT CTTAGAAC AATAGATGAA AATTTACAT TTTAAATGGA ACCCATGAG CGATGTGAT
AGCGGTGGGT GTTCAGTAA TTTTCCACT TTAATGAGA GAAATGTTA TTTTAATGTA AAATTTAC TGTGGTATCT ATGCACTACT
5201 AAATTAACTA CTGGAAATAA AACCTGGCA AGAAGTGCAGA AGACTGTTAC CGAGAACACT TCAATTTGT AAATGAGGG TTAGTGAAGA TTTAAATGGA
TTTATTAAGT GARCTTAT TACATCGGT TTCTCACTG TCTGAAATG ATGTTAAC AATCTCTCC AATACACTT AAATTACCT
5301 TGAAGATCTA AATAACCTA TAAATTGTA GAGAAATTAA TGAATGCTCA AGTAACTCA AGAACGGAA GACATACAT ATTCAAGAC TAAAGACT
ACTCTAGAT TTATGAT ATTAAACCT CTCCTTAATT ACTTACAGAT TCAATTTGT CTTGGCCTCT CGTGTGATA TAAGTACTG ATTTCTGAA
5401 ATATATGGA AGGTATACCT TCTTTTCACA TAAATTGTA GTCAATATCT TCACCCCAA AAGCTGTT GTTACCTGT CAACCTGT GAACTATC
TTATAACT TCCATATGAA AGGAAAGTGT ATTAAACAT CAGTATACA ATGTCGCTT TTTCGACAAA CATTGAACTA GTTGGAGTAA AGTTTACAT
5501 TATAGAAAGC CAAAGACAA TACACAAAT ATTCCTGTAG AACAAATGG GAAAGATGT TCCACTAAAT ATCAAGATT AGAGCAARGC ATGAGATG
ATATCTTCG GGTTCCTCT ATTGTTTAC CTTCTCA AGGTGATTA TACTCTAA TCTCTTTCG TACTCTAC
5601 TGGGGATAGA CAGTAGGCT GATAAAATAG AGTAGAGCT AGAACAGAC CCATTGATAT ATGTRAGTGA CCTPATGAAA ARATATGGCA TTTTACATG
ACCCCATCT GTCACTCGGA CTATTTCATC TCACTCTCGAG TCTTGTCTG GGTAACATA TACATTCACT GGATACCTT TTATACCGT AAATGTTAC
5701 GGAAATGAT GATCTTTTC TTTTTAGAA AACAGGGAA ATATATTAT ATGTAAGAAA TAAAGGAA CCCATATGC ATACCATACA CACAAANAAA
CTCTTACTA CTGAGAAAG AACAAATCT TTGTCCTCT TATATAATAA TACATTTT ATTTCCTT GGGTATACAS TACGTATGT GTGTTTTTT
5801 TTCCATGAA TTATAGCTT AAATGGAGAA GGCGAARCTT TAACTCTTT AGAAATAT ATAGAGCAT GCCATCATGA CTCAGTGTAA GAGAAAATT
AGGTACTT ATTATCAGA TTACCTCTT CCSTTTGAA ATTGAAAA TCTTTTATAA TATCTCTGA CGGTAGTACT GAGTCACAT CTCTTTAA
5901 TCTATGACT CAAAGCTCTA ACCACAAAGA AAAGATGTTT ATTAGATGTT CATGAAATT AAAGATTTT TTAATATAA AAATTCATTTT
AGRATCTGA GTTCAGGAT TGGTGTCTC TTCTAAACAA TAACTCTAAAC GTACTTAA TTCTGAATPA AAATTTTAT TTITGGTAA TTCTTTCTAG
6001 AGGGCCTAGA ATATTCGAA CACCCCGAGT AACGAAATTG AAATATGAG ATTATAGCAG TAAATGAG ATTATACAA AACATGCTAA AACATAAAC
TCGEGTATCT TACGTCTT ATTAAAGT GTGGGTCTT TTCTCTTACG ATTATACGT TAATTTT CTTCAGAATG TTITGTCATT TTATTTTG
6101 TAGACAAA TTGGACAGA TGAAAGGAA ACTCTTAAATA ATCAATTACAG ATGAAACT CAATCTGAA ATTCAGAGA CTATATTGC ATYATACACTA
ATCTGTCTT AACATGTCT ACTTCTCTT TGAATGTTT TGAATGTT TGAATGTT GGTAGAGCT TTAGTCTCTT GATGATRAGC TATATGTGAT
6201 AAATAGAGAA ATATTAAGG GCTAGTAAC ATCTGTGCGA ATATGGTGG TGAATGAG TGAATGAGAC AGTACTTAC CCATGGGCT
TTATCTCTT TATATTTTC CGATCTGT TGACACCGT TAACTACCC TATAATTGA ACTATACATC ACTACTTGT TCATGAAATG GTGTAACCGGA
6301 TCCCTCCCA ACCCTTACCGA CAGTAAAT CATGACAAAT ATACCTAACTT TAACTCTAA CCAGTACTCC TCAAACACTGT CAAGGTCTCAT
AGGGGGGT TGGGATGGG GTCAATTTA GTACTGTT TATGAAATT TTGGTATGG GATATGAGTT GTCAAGG AGTTTGTGCA GTTCCAGTAG
6401 AAAATTAAGA AACGCTGAG GAACTGCTAA ACTAAAGGG AACCCAAAGG GACATGAGAA TTATATGAA TGAAATGAGAT CCCAGAAACAG
TTTTTATCT TTTCGACTC TTGAGACAGTT TTGATCTCC TTGGGTCTCT CTTGACTCTT AAATACATT ACACCGTAAG ACTTACTCTA GGCTCTTGTG
6501 AAAAGAAACA GTAGCTAAA AACATATGAA ATTAATAAA AGTTGTAACCT TTAGTTTTT TAAAGAGA GTAGCTAA CACGAAAG TCATTTCAT
TTTTCTGT CATGATTTT TCAACTTGA ATCAAAATAA AAFTTTTCT CATGTTAAAT GTGCCSCTTC AGTAAARGTA
6601 ATTTCTCTG AACATTAAGT ACAACTCTAT ATTAAATAAT TTITTAATG TAGTCTGAA CATTGCCAGA AACAGAGTA CAGCGCTAT CTGCGCTTC
TAAAGAAC TGTGTTTCA TGTGAGATA TTTATGAA AAACCTTAAT TTAAATGAA ATGAGATACA TCAACGGCTT TTGTCTCAT GACAGACAG
6701 GCCTAACTAT CCAAGCTGA TTGTCTAAA ATGAGATACA TCAACGGCTT CTCATGTTT TTGTGTTTCA TGTGCTCTAT ATGAGCTT
CGGATTGATA GGTATGACT AACAGATT TACCTATGT AGTTGGAGG AGGTACAAA AACAGAGAA AAATTTCTT TTGAACTAA AAAATTCTG

Figure 7-4

TCAGAGTCCA AGTATCGTT TAACTCTCTT TCCATGTTAG TTCGACTCTT TCAAAGGGG ATAAGGATCA AATGACTCTC TAACTGTTA CTTACCCACA
 6901 TAATTTTG CAAATGCTT TTCTGTGTTT ATCAATANGA CCATGTGATT TTCTCTTA ACCGTGTGAT GGGACARATT AGCTTAATTC ATTTCAAC
 ATTAACAA GTTACGAA AGACACAGA TTGTTACTT GGTACACTA AAGGAAAT TGGACAACTA CCTGTGTTA TGCATTAAC TAAAGTTG
 7001 GTGAAACC CCTACATCTT CTGGATAAA TCTACTTGG TTGTGGTA TATTTTGA TACATTCTG GATTCCTTTT GCTATAATTG TTGTGAAAT
 CRACTGGTG GGAACTT GACCTTATTT AAGATGAAAC AACCAACAT ATAAACACT ATGTTAGAAC CTAAAGAAA CGATTAAAC ACHACTTTA
 7101 GTTGTATCT TTGTTCTGAGATATTG TCTGTGTTT TCTTTCTG TAATGCTATT TTCTAGTCTC GTTATTTGG TAATGCTGG CTAGTGTAT
 CAAACATAGA AACAGTACT CTCATAAC AGAACAAAGC ATTPACAGTA ARGATGAAAC ATTACGACG CCTAAATTC ATTACGACG GATGACTTA
 7201 GATTTAGGA GATATCCCTC TGCTTCTGTC TTCTGAGTA CGCGGGGGC CGCTGTTT ACAGCTGCT GATCTGGAA ACCCTGGGT TACCGAACTT
 CTAAATCCCTT CATAAGGGAG AGACTCTCAT GGGACGAGA AGACTCTCAT GGCAGCRAAA TTGTGAGCA CTGACCCCTT TGGACCCGA ATGGGTGAA
 7301 ATGGCCCTTG CAGCACATCC CCCCTTCGTC AGCTGGGTTA ATAGGAAGA GGCGCGCACC GATGCGCCRTT CCAACAGTTT GCGCAGGCCG AATGGGAAAT
 TTACGGAAAC GTCGTGTAGG GGGAAAGGGG TOGACCGAT TATGCTCTC CGCGGGGG CTAGGGAA GGTTGTCAA CGCGTGCAC TTACCGCTA
 7401 GGGCCTGTAT GCGGTATTTT CTCCTACGC ATCTGTGGG TATTCACGC CGCATACGC ARAGGAACCA TAGTACGCC CCTGTAGGG CGCATTAAGC
 CGCGGACTA CGCCATAAA GAGGATGCG TRGACACCC ATAAAGTGT GCGTGTGCG TTTCGTTGT ATCATGCGC CGACATCSCC GCGTAATTGCG
 7501 GGGGGGGG TGGTGTATAC CGCGAGGGT ACCGCTACAC TTGCGAGGCC CCTAGGCC GCTAGCTTCC CTTCTCTTC GCACTTCTC
 CGCGGCCAAC ACCACCATG CGCGTGCAC TGGCGATG AACGGTGCAG GGATGCGGG CGAGGAAAGC GAGAGAGGG FAGGAAGAG CGGTGCAAGC
 7601 CGGCTTCTC CCGTCAGCT CTAATCGGG GGCTCCCTT AGGGTTGGA TTATGTCCT TATGCCACCT CGACCCCAA AACITGATTT TGGGTGTTG
 GCGGAAAGG GCGATGTCGA GATTPGCCC CGAGATGCGA AAATCAGAA ATGCGTGGG GCTGGGTTT TTGAACTAA ACCGACTTAC
 7701 TTACAGTAGT GGGCCTATCGC CCTGATGAC GGTGTTTCGC CCTTTGACTG TTGGTCCAC GTTCTTTAAT AGTGGACTCT TGTCCAAAC
 ATGTCATCA CCCGTRGCG GGACTATCG CCAAAGCG GAAACTGCA ACCTTGTG CRAGAAATTA TCACTTGAGA ACAGGTTG ACCTTGTG
 7801 CTACACCTA TCTCGGCTA TTCTTTGAT TTAAAGGA TTGTGCGAT TTGGCCAT ACCGAGATA TACTGACTA ATTACAAATAA TTAAACAAA TTAAACGCGA
 GAGTTGGGAT AGAGCCGAT AAGAAACATA ATATTCCTT AAACGGCTA AAGCGGATA ACCGAGATA TACTGACTA ATTACAAATAA TTAAACAAA TTAAACGCGA
 7901 ATTTAACAA ATATTAACG TTACATTAT TATGGTGCAC TCTCAGTACA ATCTGCTCTG ATGGCGATA GTAAAGCCG CCGCGACACC CGCAACAC
 TAAATTTGTT TTATTAACG AAATGTTAAATACCGAT AGTACGATG TAACTGAGAC TACCGCGTAT CAACTGGTC GGGCTGTGG CGGTTGTTGCGT
 8001 CGGTGACGG CCCTGCGG CCTGTCATCC GCTTACAGAC AAGGTGTC CGTCTCGGG AGCTGATGT GTCAAGAGTT TCACTGCGT
 GCGACTGCG GGGACTGCCC GAACAGRCGA GGGCGCTAGG CGATGTCCTG TTGCACTG GCAGGGCCC TCACTGCTCAA AAGTGGCAAGT
 8101 TCACCGAAC GGGCAGAGA CGAAAGGCC CGTGTGATAG CCTATTATAAT TAGTATAAT ATGGTTCT TAGACGTGAG GNGGCACTT
 AGTGCTTG CGCGCTCTCTC GCTTTCGGG AGCCTATGC GGTAAARAT ATCCATAC AGTACTTAA TCACTRAGA ATCTGCGTGC CACCGTGAAGA
 8201 TCGGGAAAT GTGCGGGAA CCCTTATTTG TTATTTTC TAATACATT CAATATGA TCCGTCTG AGACATAAC CCTGATAAT GCTTCATAA
 AGGCCCTTCA CACGGCCCTT GGGGTAAC AAATTAAG ATTATGTA GTTATACAT AGGGGAGTAC TCTGTATATG GAGCTATTTA CGAAGTTAT
 8301 TTGTAAAAA GGAAGAGTAT GAGTATCAA CATTTCGGTG TGCCCTTAT TCCCTTTT GCGCATTTT GCTTCCTGT TTTGGTCAC CCAGAACGC
 ATTAACCTTT CTCCTCTATA CTCATAGT GTAAAGGCAC AGGGGATA AGGGAAAAA CGCGTAAAAA CGGAGGACA AAAACGAGTG GCTTCATAA
 8401 TTGTGAAAGT AAGATGCTT GAGATCAGT TGGGTGCAAG AGTGGGTAC ATCGAATCTG ATCTCAGACAG CGGTAGTGT TTOGCCCCUAA
 ACCACATTCA TTTCCTACGA CTTCTPAGTC ACCCAGCTGC TCACTCCTG TAGCTGAC TCACTCTAG GAACCTCTAA AAGCGGGGTT
 8501 AGAACGTTT CCAATGATGA GCACCTTAA AGTTCGCTA TGTGGGGG TATTATCCCG TATGACGCC GGGCAAGAGC AACCTGGTC CGGCAVACAC
 TCTTGCAAAA GTTACTACT CGTGAATAAT TCAAGACGAT ACACCGCCG ATAATAGGGC ATAPCTGCGG CCCTGTTCTG TTGRCGACG GCGGTATG

Figure 7-5

8601 TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGGATCT TAGGGATGGC ATGACAGTAA GAGGATTAG CAGTGGCTGCC ATMACCAAGA ATAGAGTCT TACTGACCA ACCTGAGTGG ATGCTCTAGA ATGCTACCGG TACTGCTATT CTCCTTATT

8701 GIGATAACAC TGGGCCAAC TTACTCTGA CAGGATCGG AGGACCGAAG GAGCTAACCG CTTTTTSGCA CAACTGGGG GATCATGGTA CTCGCCTTG CACTATGTT AGCCGGGTG ATGAGACT GTTGTAGCC TCCGGCTTC GAAAGAAGT GTGTAACCC CTTAGTACATT GAGGGAACT

8801 TGGTGGAA CGGAGGCTGA ATGAGGCAAT ACCAACGAC (GAGCGTACA CCACGATGCC TGTAAGCTG GCAACAACTG TGCGCAACT ATTAACTGCC AGCAACCTT GGCCTCGACT 'TACTTCGGTA TGGTTGGCTG CTCGCACTGT GGTGTCAGG ACATCGTAC CTTGTGGA TAATTGACCG

8901 GAACTACTTA CTCTAGCTTC CGGGCAACCA TTAACTGACT GGATGGGGC GGATTAAGT GGAGGACCC TTGTGGCTG CCGCCCTGG GCAACGGACCA CTGTGAGTAT GAGATCGAG GGCCTGGTGT CCTATCTCG CCTATTCGA CGTCTGGTGA AGACGGAG

9001 TTATTGCTGA TAAATCTGGA GCGGTTGACG GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCAGATGG TAAGGCCCTCC CGTATCGTAG TTATCTCAC AGTATGCTAC ATTAACGACT ATTTAGACCT CGTCTGACCC CGCGTCTAAC ATTICGGGAGG GCAATGCTAC AATAGATGT

9101 GACGGGGAGT CAGGCAACTA TGGATGACG AAATAGACAG ATTCGCTGAGA TAGGCTGAGA ACTGATTAAG CATGGTAAAC TGTCAGACCA AGTTTACCA CGGCCCTCA GTCCGTTGAT ACCTACTTC TTATCTGTC TAGCCTACTT ATCCACGGG TGACTATTACG ACAGTCGGT TCAATGTT

9201 TATATACTT AGATGATTT AAACACTTAT TTAAATTTA ARAAGATCTA GGTTGAGATC CTTTGTATA ATCTCATGAC CAAATCCCT TAACGTGAGT ATTTATGAA TTCTRACTAA TTTCCTGAGT CCACCTCTAG GAAATCTAT TAAAGTACTG GTFPTAGGG ATTGCACTA

9301 TTTCGTTCCA CTGAGGCTCA GACCCCGTAGA AAGGATCAA AGGATCTCTT TGAGATCCTT TTTCCTGCG CGTAAATGCC 'TGCCTGCCAA CAAAAACCC ATGGAGGT GACTCGCTGT CTGGGGCTTC TTTCCTGAGT CCTCTGAGA ACTCTAGGA AAAAAGAGGC GCAATGAGC AGCAGCTT GTTTTTTGG

9401 ACGGTACCA CGGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCACACTTTT TTCCGGAGGT AACCTGGCTTC AGCAGGGCC AGATACCAA TACTGTTCTT TGGCTGGT CGCCACCAA AAGCTGGCT AGTCTCGAT GGTTGAGAAA TGGCTTCCA TTGACCGAAG TGCTCTCGG TCTATGGTT ATGACAAAGA

9501 CTAGTSTAGC CGTAGTCTAGG CCACCACTCTG AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCTGTAAAC AGTGGCTGCTG GCCAGTGGCG GATCACATCG GATCATCG GATCATCG GETGGTGAAG TTCTTGAAC ATCGTGGGG ATGTTGGAG CGAGACATT AGACAATGG TCAACGAGCA CGGTCAACGGC

9601 ATAAAGTGTG TCTTACGGG TTGGACTCA GAGGATAGT ACCGGATAAG GCGCGAGGT CGGGCTGAAAC GGGGGTCTGG 'TGCACACGG CCAGCTTGGAA TATTCTGAC AGAATGCC AACCTGAGT CTGCTATCA TGGCTATTCC CGCCGTCGCA GCCGCACTTG CCCCCTCAGG ACCTGTTGCG GTGCAACTC

9701 GCGAAAGGCC TACACCGAAC TGAGATCTC AGINGCTGAG CTATGAGAA GCGCAGCGGC TCCGGAGGG AGAAAGGGG AGCAGTATC GGTAAAGCGC CGCTGTCTGG ATGTGGCTTG ATCTCTAGGA TCTGCGACTC GATRACTCTT CGCGCTGCC TCTTCTGCC 'TSTCCCATGCG CCATTCGCG

9801 AGGGCTGGAA CAGGAGRGCG CACGAGGGG CTTCAGGGG GAAACGCTG GTATTTTAT AGCTCTGTCG GGTTCGCC CCTCTGACTT GAGGGTCTGAT TCCCGCCCTT GTCCCTCTGC GAGGTCTCC CTTGCGGAC CATAGAATAA TCAAGACGC AGAGCTGAA CTCGCACTA

9901 TTITSTGATG CTCGTCAGGG GGGCGGAGAC TATGGAAAAGA CGCGAGCAAC GCGGCTTTTACGGTCTCTT GGGCTTTCG TGGCTTTCG TGCCTCTTG CTCACATGAA AAGAACTAC GAGGAGTCCC CCGGCCCTGG ATACCTTTT GCGGCTGTTG CGCCGAAAATGCAAGGA CGGGAAAAC GAGGTTACAA

10001 CTTCTCTGCG TTATCCCTG ATTCCTGGA TAACCTGTT ACCGGTTTGG AGTGGTGA TACCGCTGC CGCAGGGAA CGACGGAGG CAGGGTCTGC GTCGCTGCTT GCTGGCTCGC

10101 GTAGGGAGG AGGGGAGA GCGGGGGC AAGGTCGCA CGCACAGAT CAACTTACG AATGAGTACG TCTCTTCTG CAAATGAG GTGAAATTT₄ CACTCGCTCC TTTCGCTCTT CCGGCCCTGG TCCACCGGT GCTGTCTCA GTTATTTTC GAGGAAATC GTTATTTTAA

10201 ACATTTRAA ATGAGACCA TAGAGCAGTGT ATGAAATAA TCTACTGG AATAAACTA GCGCAAGAGG TGCAAGACTG TACCTACAA TGAAATGGT GTAAATTT TTACTGTTG ATCTGCTAC TACTTTATT AGTGAACCT TTATTTAGAT CGGTCTCTTC AGCTGCTGAC ATGGCTCTT

10301 TTGTAATGA GAGGTTAGTG AAGATTTAA TGTATGAGA TCTAAATAA CTATTAATT GTGGAGAAA TTAATGATG TCTAAGTTAA TGCAGAAACG AACATTTACT CTCCATCAC TTCTARATT ACTACTCTT AGATTTATT GAAATTTAA CACTCTCTT AATTTAA

Figure 7-6

1 10401 GAGGACATA CTATTCAT GAACTTAAAG ACTTCTTT CACATAATT TGTCATCAAT ATGTCACCC CAAAAAGCT CTCCTCTAT GATTTAGTA CTCCTCCATA CTCCTCCATA TGAAAGAA TGTTTTCGCA
10501 GTTGTTAAC TTGTCACCT CATTGCAAA ACAATAACCA AAATATCTT GTAGACAAAT ATGGAAAGA ATGTTCCACT CAAACAAATG ARCACTGGG TTAAAGTTT ACATATATCTT TCTGTTTCTC TCAAGGTGA
10601 AAATATCAAG ATTAGAGCA AAGCATGAGA TGTGTTGGGA TAGACAGTGA GGCTGTATAA ATAGAGTGA GCTCGAAAC AGRCCTATG ATATATGTTAATGTTTACCTT TCAAGGTGA
10701 GTGACCTATG AAAAATAT GGCACTTAC ATGGGAAA TGATGATCT TTCTCTTTT AGGAAACAG GGAATATAT TTATATGTTAA AAATAAAAG CACTGATAC TTTCCTTAA CGTCTCTC ATCTGTCATC ACACCCCT
10801 GGAACCCATA TGATATACCA TACACCAA AAAATCCAG TGAACTATAA GTCTAAATGG AGRAAGCAA ACITTAATC TTTCAGAAA TTAATAGAA CCTTGGTAT ACATATGGT ATGTTGGT TTAACTTCTT ATTATATCTT
10901 GCGTGCCTAC ATGCTTCAG TGATGAGAA ATTCTCTT GACTCAAGT CCAAACTCA AGAAAGAT TGTTAATGG ATTGATGAA TATTAAGACT CGPRGGTAG TACGTGAGC ACATCTCTT TTAACTGATA
11001 TTTTTAAATTAATAC CATTAGRAA TGTCAGGCCA TGRAATGCA GAAATATTGTTT GCAACCCCT AGRAAGAGA ATTGTATAT GCGATATAA ATTAATTTTAAATTAATAC CTTTATATAA CGTTGIGGG TCATCTCT TCACTTATTA
11101 AAAGAAAGTC TTACAAATCA GTAAAGATAA AACTAGRAA AAATTTGAA CGATGAAAG AGAAACTCTA AAATATCATT ACATGAGA BACTCACTT TTTCCTTCAG ATGTTTGTAT CTTTAACTT TTGTTGAGT TTATAGTAA TGTTGACTCT TTGAGTTAGA
11201 CAGRATCAG AGACTATCA TTGATATAC ACTAATTTAG AGAAATATAA ARAGGTAAAG TAACATCTG GGCATATG ATGGATATAA ACCTGATAT GTCCTTACGCTC ATGTTGAGT TTTCGAGTC ATGTTTAAAC TCCCTTAAAC TCACTTATTA
11301 GATGTTGATCA GAACTGACT TTACCCATG GGCTTCCTCC CCAACCCCT ACCCCAGTAA ATATCAGAC AATTAATCTT TAAACCAT TACCCPATAAT CTACACTCT CTTGTCATGA ATGGGGTAC CGGARGGGG
11401 CTACACATA CTTCACAA CTGTCAGGT CTCAGGACTG TCAAAATPAA GAGGAACCC AGGAGACATG AGGATTTAT GATGGTGTAT GAGGTTTCA GTAGTTTTA
11501 GTAATGSGC ATTCTGAATG AGATCCCGAGA ACAGAAAGB AACAGTAGCT ABAAACTAA TGRAATATAA ATAAAGTTTG AACITTAGTT TTCTCTGGGT TCCTCTGTAC TCCTTAATTA
11601 AAGAGTACCA TPAACCGGC AAAGTCATT TCAATTTT CTGATGCTC TTCTCTTC TTGTCATGAA TTTTGTATT ACTTATTT TATTCAAC TTGAACTCAA AAATAATTTC
11701 CAGRAACGCA AGTACAGCGAG CTATCCATG TGTCGCCTAA CTATCCATG CTGATGGTC TAAATGAGA TACATCAAG CTCCCTCCATG TTTCGGTTT
11801 TCTTTTAAATGAAACTT TATTTTTAA GAGGAGTTG AGGTTCATAG CAAATGAG AGGAAGGTAC ATTCAGCTG AGGAGTTT TTCTCTTCAGC TCCCTCAAA GGAGATAAGG
11901 TAGTTTACTG AGRGTTGCA TGCTGATGG GTGTTAAATT TTGTCATGAT ATGACATGT GATTTCCTC TCTCTATCTG GAGGGGTAC AAAAAGCTC
ATCAAAATGAC TCTCTAACTG AGTACTTAC CACATTTAA ARACGTTAAC CAGATGTTA TACTGTTACA CTAAGGAGA AAATTGGACA
12001 TGATGGGCRCA ATTCAGCTA ATATCTGGAA TAAATTCAC AGACGTTGAA CCACCCCTAC TCTCTATCTG TTGTTGNGG TCTATATT TTGATACATT
ACTACCCCTGT TTATGCTAT TACATTAAGG TTGCAACTT GGTGGGAATG TATAGCTT ATTAGATG ACCAACACCC ACATATAAA AACATGTTAA

Figure 7-7

GACCCCTAGA AAGCGGTTA TAAACACACT TTACACAA TAGAACAG AACCTCTAT AACAGACAA CAAAGAAA GACATTAA GTAAAGATC
 12201 TTCCGGTATT AGGTAATGC TGGCCTAGT GAATGATTAA GGAGTATT CCTCTGCTTC TGTCCTG TGTCCTGCTTC CGCCAATAC GCAAACGCC
 AGGGCTAA TTCCATTACG ACCGGATCA CTACTAATC CCTICATAAG GGAGCGAAG ACAGAAGCT CGGGTTTANG CGTTGGGG
 12301 TCTCCCGCGA CGTGGCGA TICATTAACT CAGCTGGCAG CAGCTGGCA TICATTAACT CCGAGTTTC CCGAGCTGGAA AGCGGGCAGT GAGTAGCTC
 AGAGGGCGC GCAACCGCT AAGTAATAC GTGACCGTG CTGACCGTG CTGACCGTG CGCTGACCTT TCGCCCGTCA CTGCGGTGCTC CTCATCGAC
 12401 ACTCATTTAGG CACCCAGG TTACACIT TTACACIT ATGCTTCGG CTGCTATGTT GTGTCGANT GTGAGCGGAT AACAATTCA CACAGGAAC AGCTATGCA
 TGAGTATCC GTGGGGTCCG AATGTGAA TAGTGGAA CACACCTAA CACATACAA CACTCGCTA TGTGTTAAGT GTSTCCCTTG TCGATAGT
 12501 TGATPACGAA TTAA
 ACTATGCTT ATT
 >length: 12514

Figure 7-8

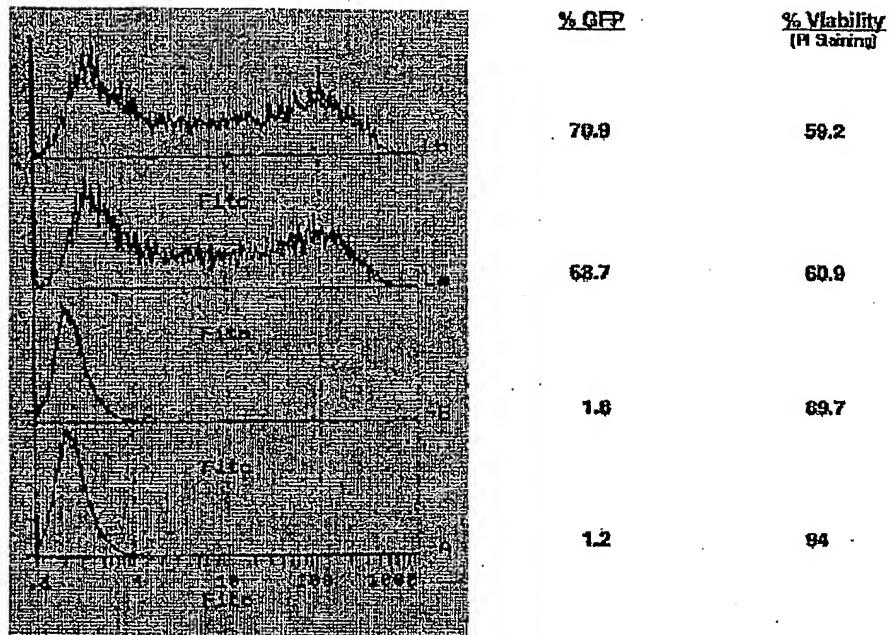
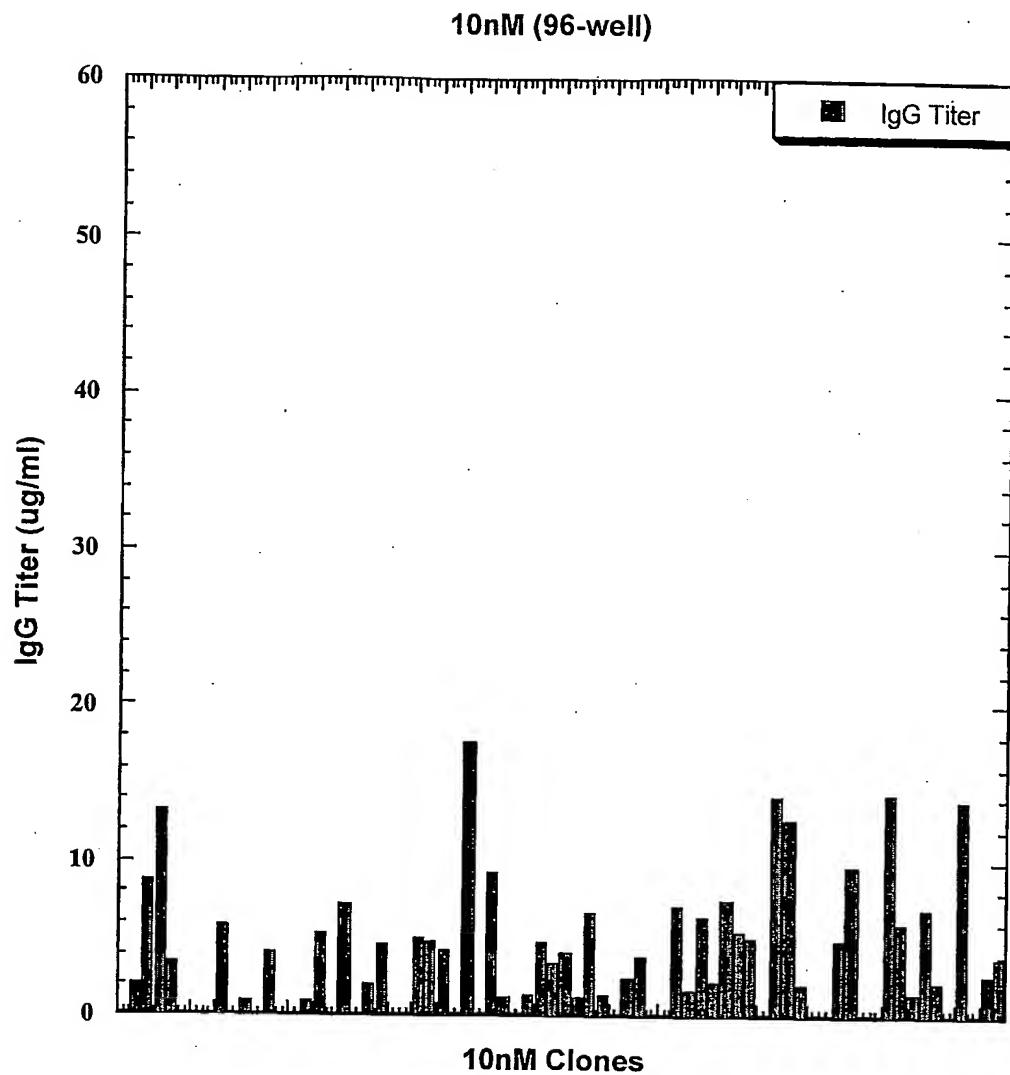


Figure 8. FACS analysis of transiently transfected CHO cells with a GFP plasmid in 250 ml spinner transfection.



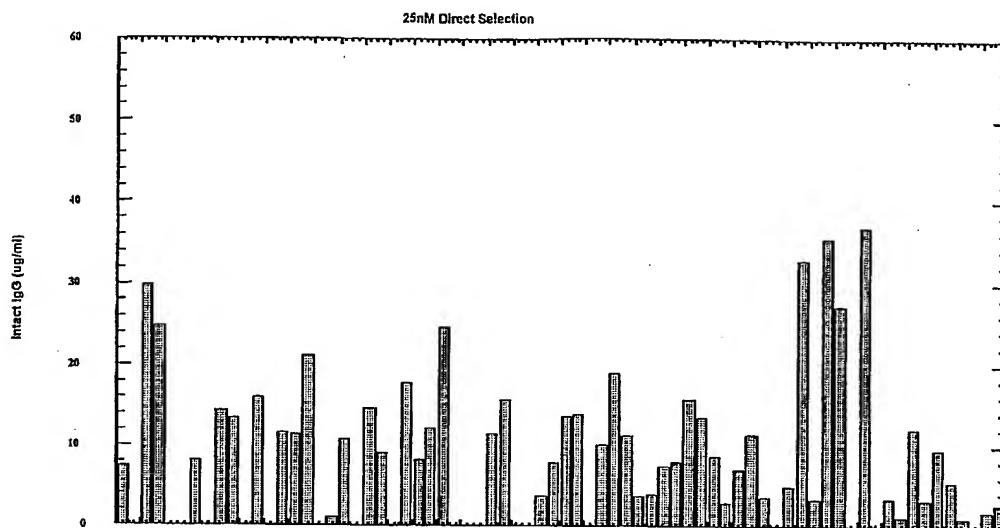


Figure 10-1

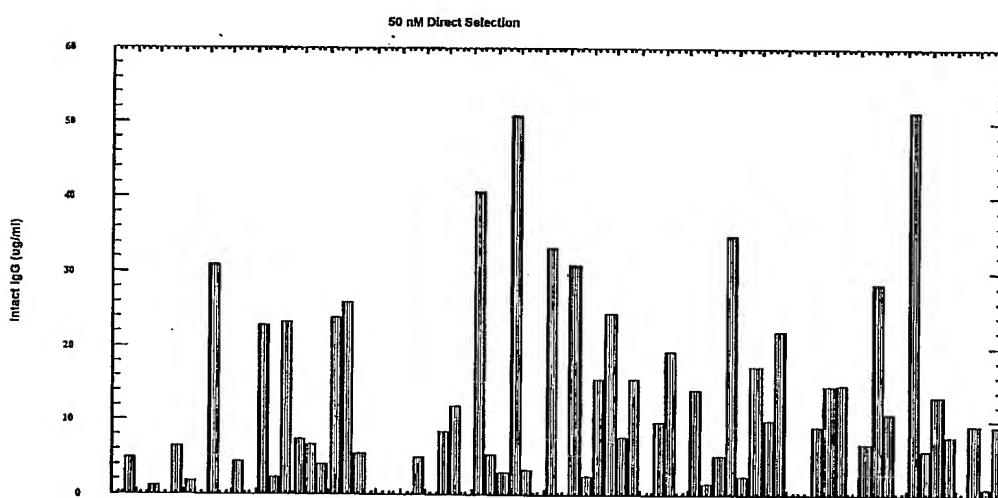


Figure 10-2

Figures 10.1 and 10.2. Expression level of clones from 25 and 50 nM MTX direct selections of SV40-based constructs derived from spinner transfection, respectively.

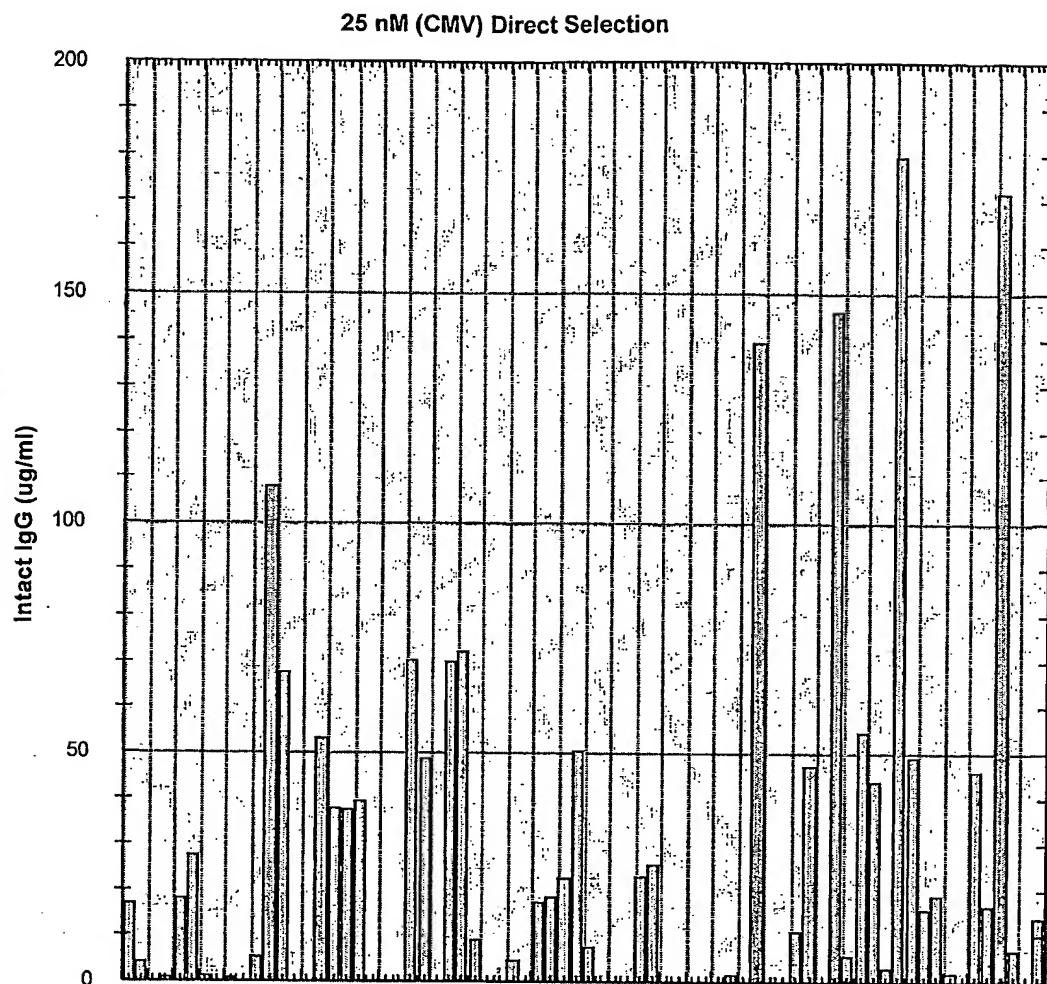


Figure 11. Expression level of clones from 25 nM MTX direct selection of CMV construct derived from spinner transfection.

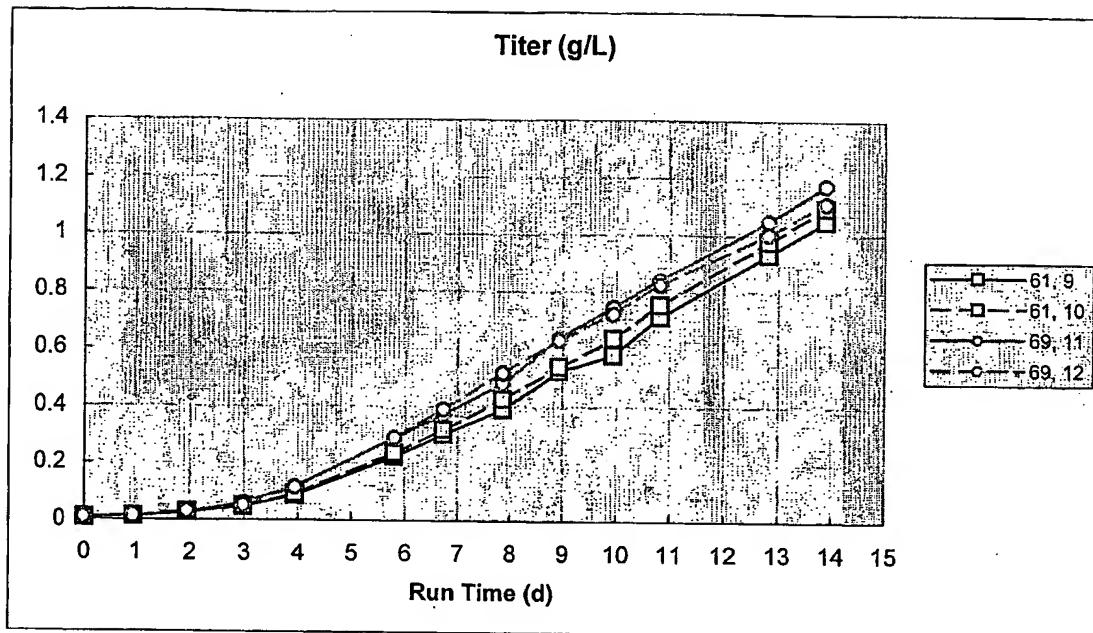


Figure 12. Titer Evaluation in Miniferm.

Figure 13. Plasmid pCMV.IPD.Heterologous Polypeptide

5 <400>
 60 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC ACCGGTAGTA ATCAAATTACG
 120 GGGTCATTAG TTCACTAGCCC ATATATGGAG TTCCGGCTTA CATAACTTAC GGTAATGGC
 180 CGGCCTGGCT GACCGCCCAA CGACCCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC
 240 ATAGTAACGC CAATAAGGGAC TTTCATTGAG CGTCAATGGG TGGAGTATT ACGGTAAGCT
 300 GCCCCTACTGG CAGTACATCA AGGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT
 360 GACGGTAAAT GGCCCCGCTG GCATTATGCC CAGTACATGA CCTTATGGGA CTTTCCTACT
 420 TGGCAGTACA TCTACGGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC
 480 ATCAATGGC GTGGATAGGC GTTGTACTCA CGGGGATTC CAAGTCTCCA CCCCATGGAC
 540 GTCAATGGG GTTGTGTTTG GCACCAAAAT CAACGGACT TTCCAAAATG TCGTAACAC
 600 TCCGGCCCAT TGACGCCAAAT GGCGGTAGG CGTGTACGGT GGGAGGCTA TATAAGCAGA
 660 GCTCGTTAG TGAACCGTCA GATGCCCTGG AGACGCCATC CACGGCTTT TGACCTGGC
 720 CGGGCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGMAGT AGTGGAGGG CTTTTTGGA
 780 GGCCTAGGCT TTTGCAAAAAA GCTAGCTTAT CGGGCGGGG ACGGTGCATT GGAAACGGGA
 840 TTCCCCGTGC CAAGAGTGCAC GTAAAGTACCG CCTATAGAGC GACTAGTCCA CCATGACCGA
 900 GTACAAGCCC ACGGTGCGCC TCGCCACCCG CGACGAGTC CGGGGGCCG TACGGCACCT

Figure 13.1

960 CGCCGGCGG TTGGCCGACT ACCCGCCAC GCGCCACACC GTAGACCCGG ACCGCCACAT
 1020 CGAGGGTTC ACCGAGCTGC AAGAACTCTT CCTCACGCCG GTCGGGCTCG ACATGGCAA
 1080 GGTGGGGTC GCGGAGCCGACG GCGCCGGGT GCGGGTCTGG ACCACGCCGG AGAGGGTGA
 1140 AGCGGGCGG GTGTTGCGCG AGATGGCCC GGCATGGCC GAGTTGAGCG GTTCCCGGCT
 1200 GCGCGCGAG CAACAGATGG AAGGCTCCCT GGGCGCAC CGGGCAAGG AGCCCGGTG
 1260 GTTCCTGGCC ACCGTCGGGG TCTCGCCCGA CCACCAGGGC AAGGGTCTGG GCAGGGCGT
 1320 CGTGCTCCCC GGAGTGGAGG CGGGCGAGCG CGGGGGTGC CCCGCCCTTC TGGAGACCTC
 1380 CGCGCCGGC AACCTCCCCCT TCTACGAGCG GCTCGGCTTC ACCGGTACCCG .CCGACGTGGA
 1440 GGTGCCGAA GGACCCGGCA CCTGGTGCAT GACCGGCAAG CCCGGTGCCA ACATGGTTG
 1500 ACCATTGAAAC TGGCATCGTC CGCGTGTCCCA AAATATGGGG ATTGGAAAGA ACGGAGACT
 1560 ACCCTGGCCT CCGCTCAGGA ACGGTTCAA GTACTTCAA AGAATGACCA CAACCTCTTC
 1620 AGTGGAAAGTT AACAGAAATC TGGGTGATTAT GGCTAGGGAA ACCTGGTTCT CCATTCCTGA
 1680 GAAGAATCGA CCTTTAAAGG ACAGAAATAA TATAGTCTC AGTAGAGAAC TCAAAGAAC
 1740 ACCACGAGGA GCTCATTTTC TGGCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGGCAGTT CTGTTTACCA
 1800 ACCGGAATTG GCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGGCAGTT CTGTTTACCA
 1860 GGAAGCCATG AATCAACCAAG GCCACCTCAG ACTCTTGTG ACAAGGATCA TGCAGGAATT
 1920 TGAAAGTGCAC ACGTTCCTCC CAGAAATTGA TTGGGGAA TATTAACCTC TCCCAAGAATA
 1980 CCCAGGGTTC CTCTCTGAGG TCCAGGAGGA AAAAGGATC AAGTATAAGT TTGAAGTCTA

2040 CGAGAAAGAAA GACTAACGTT AACTGCTCCC CTCCTAAAGC TATGCATTIT TATAAGACCA
 2100 TGAGACTTTT GCTGGCTTTA GATCCCCCTTG GCTTCGTTAG AACCGAGCTA CAATTAAATAC
 2160 ATAACCTTAT GTATCATACA CATACTGATT AGGTGACACT ATGAATAAC ATCCACITRG
 2220 CCTTTCTCTC CACAGGTGTC CACTCCAGG TCCAAC TGCA CCTCGTTCT ATCGATTGAA

TTCCACC --Insert Sequence of Interest--

CGA TGGCCGCCAT GGCCCAACT GTTTATTGCA GCTTATAATG
 GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTT TCACTGCATT
 CTAGTTGTTG TTTGTCAAA CTCATCAATG TATCTTATCA TGTCCTGGATC GGGAAATTAT
 TCGGGCGAGC ACCATGGCCT GAAATAACCT CTGAAAGAGG AACTTGGTTA GGTAACCTATT
 AATAGTAATC AATTACGGGG TCATTAGTTC ATAGCCATA TATGGAGTTC CGCGTTACAT
 AACTTACGGT AAATGGCCCG CCTGGCTGAC CGCCCAACGA CCCCGCCCA TTGACGTOAA
 TAATGACGTA TGTTCCCATA GTAAQGCCAA TAGGGACTTT CCATTGACGT CAATGGGTGG
 AGTATTAGC GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC
 CCCCTATTGA CGTCAATTGAC GGTAAATGGC CCGCCTGGCA TTATGCCAG TACATGACT
 TATGGGACTT TCCTACTTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTA
 TGCCTTTG GCAGTACATC AATGGGGGTG GATAGCGGT TGACTCACGG GGATTTCAC
 GTCTCCACCC CATTGACGTC AATGGGAGTT TGTTTGCA CCAAAATCAA CGGGACTTTC
 CAAAATGTCG TAACAACTCC GCCCCATTGA CGCAAATGGG CGTAGGCCT GTACGGTGGG

AGGTCTATAT AAGCAGAGCT CGTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC
GCTGTTTGA CCTGCTAGCT TATCCGGCCG GAAACGGTGC ATTGAAACGC GGATCCCCCG
TGCCCAAGAGT CAGGTAGTA CGCCTATAG AGTCTATAGG CCCACCCCT TGCTTCGTT
AGAACGGGC TACAATTAAAT ACATAACCTT TTGGATCGAT CCTACTGACA CTGACATCCA
CTTTTCTTTT TTCTCACAG GTGTCCACTC CCAGGTCCAA CTGACACTCG GTTCGCGAAG
CTCGCTTGGG CTGCATCGAT TGATTCCAC C --Insert Sequence of Interest--
CGATGG CGGCCATGG CCAACTTGT TATGCAAGCT TATAATGTT
ACAAATAAG CAATAGCATC ACAATTCA CAAATAAGC ATTTCATTCATCTCA CTGCATTCTA
GTTGTTGGTT GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGG ATTAAATTG
GGCAGGCACC ATGGCTGAA ATAAGTTAA ACCCTCTGAA AGAGGAACCT GTTAGGTAC
CGACTAGTCT TTGCAAAAAA GTCTTACCT CGAGGGCCG CTTAATTAAG GCGGCCATT
TAAATCCTGC AGGTAACAGC TTGGCACTGG CCGTCGTTT ACAACGTGTT GACTGGAAA
ACCCTGGCT TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTCGCC AGCTGGGCTA
ATAGCGAAGA GGCGCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGGAAAT
GGCGCCTGAT GGGTATTT CTCCCTACGC ATCTGTGCGG TATTCACAC CGCATACGTC
AAAGCAACCA TAGTACGCGC CCTGTAGCGG CGCATTAAGC GCGGGGGTG TGGTGGGTAC
GGGAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGGCCG GCTCCTTCG CTTTCCTCCC
TTCCCTTCCTC GCCACGTTGC CGGCTTTC CCGTCAAGCT CTAATCGGG GGCTCCCTT

AGGGTTCCGA TTAGGTGCTT TACGGCACCT CGACCCCCAA AACATGATT TGGGTGATGG
TTCACGTAGT GGGCCATCGC CCTGTAGAC GGTGTTTCGC CCTTTGACGT TGGAGTCAC
GTTCTTTAAT AGTGGACTCT TTGTTCCAAAC TGGAAACAA CA CTCACCCCTA TCTCGGGCTA
TTCTTTGAT TTATAAGGG A TTGTCGCGAT TTGGCCCTAT TGGTTAAAAA ATGAGCTGAT
TTAACAAAAA TTAAACGCGA ATTAAACAA ATATTAACG TTACAAATT TATGGTGCAC
TCTCAGTACA ATCTGGCTCTG ATGCCGCATA GTTAAGCCAG CCCCGACACC GCCCCGACAC
CGGCCAACAC CGGCTGACGC GCCCTGACGG GCTTGTCTGC TCCCGGCATC CGCTTACAGA
CAAGCTGTGA CGGCTCTCCGG GAGCTGCATG TGTCAAGGGT TTTCACCGTC ATCACCGAAA
CGCGCAGAG AGCAAAAGGGC CTCGTGATAAC GCCTTATTATTT ATAGGTTAAT GTCATGATAA
TAATGGTTTC TTAGACGTCA GGTGGCACTT TTGGGGAAA TGTGGGGAA ACCCCTATT
GTTTATTTTT CTAATACAT TCAAATATGT ATCCGGCTCAT GAGACAATAA CCCTGTATAAA
TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATCA ACATTTCGGT GTGCCCTTA
TTCCTTTT TGCGGCAATT TGCGTTCCGT TTTTGCTCA CCCAGAAACG CTGGTGAAG
TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGTTA CATCGAACTG GATCTCAACA
GCGGTAAGAT CCTTGAGAGT TTGCCCCCG AAGAACGTT TCCATATGATG AGCACTTTA
AAGTTCTGCT ATGGGGCGC GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTGGTC
GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATT
TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGTGTC CATAACCATG AGTGTATAACA

CTGGGGCAA CTTACTTCTG ACAAAGATCG GAGGACCGAA GGAGCTAAC GCTTTTTGCA
ACAAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGG ACCGGAGCTG AATGAAGCCA
TACCAAAAGGA CGAGCGGTGAC ACCACGATGC CTGTAAGCAAT GGCAACAACG TTGGCGAAAC
TATTAACTGG CGAAACTACTT ACTCTAGCTT CCCGGCAACA ATTAAATAGAC TGGATGGAGG
GGATAAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCGGCTGG TTTATTGCTG
ATAAATCTGG AGCCGGGTGAG CGGGTATCAT TGCAAGCACTG GGGCCAGATG
GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC
GAATAGACCA GATCGCTGAG ATAGGTGCT CACTGATTAA GCATTGGTAA CTGTCAGACC
AAGTTTACTC ATATATACTT TAGATTGATT TAAAACCTCA TTTTAATTAAAGGATCT
AGGTGAAGAT CCTTTTGAT AATCTCATGA CCAAATCCC TTAACGTGAG TTTTCGTTCC
ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCTT TTTTTCTGC
GGGTAATCTG CTGCTGCAA AAAAAAAAC CACCGCTACC AGCGGTGTT TGTTCGCGG
ATCAAGAGCT ACCAAACTCTT TTCCGAAAGG TAACTGGCTT CAGCAGAGCG CAGATACCA
ATACTGTTCT TCTAGTGTAG CGTAGTTAG GCCACCACTT CAAGAACTCT GTAGGCACCGC
CTACATACCT CGCTCTGCTA ATCTGTAC CAGTGGCTGC TGCGAGTGGC GATAAGTCGT
GTCCTACCGG GTGGACTCA AGAGATAGT TACCGGATAA GGCGCAGGG TCGGGCTGAA
CGGGGGTTC GTGCACACAG CCCAGCTGG AGCGAACGAC CTACACCGAA CTGAGATACC
TACAGCGTGA GCTATGAGAA AGGCCACGC TTCCCGAAGG GAGAAAGGG GACAGGGTATC

CGGTAAGCGG CAGGGTCGGA ACAGGAGGC GCACGGAGGA GCTTCCAGGG GGAAACGGCT
GGTATCTTAA TAGTCTCTGTC GGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTGTGAT
GCTCGTCAGG GGGGGGGAGC CTATGGAAA ACGCCAGCAA CGGGGCCCTT TTACGGTTCC
TGGCCTTTTG CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCTT GATTCTGTGG
ATAACCGTAT TACCGCCCTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGGAAACCG CCTCTCCCG
CGCGTTGGCC GATTCAATTAA TGGAGCTGGC ACGACAGGTT TCCCGACTGG AAAGGGCA
GTGAGCGCAA CGCAATTAAAT GTGAGTTAGC TCACTCATA GGCAACCCAG 'GCTTTACACT
TTATGCTTCG GGCTCGTATG TTGTGTGGAA TTGTGACCGG ATAACAATT CACACAGGAA
ACAGCTATGA CATGATTACG AATTAA

Figure 14. Plasmid SV40.NPD.Heterologous Polypeptide

6	<400>
60	TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT
120	CAGTTAGGGT GTGAAAGTC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT
180	CTCAATTAGT CAGCAACCAG GTGTGAAAG TCCCCAGGCT CCCCAAGCAGG CAGAAAGTATG
240	CAAAGCATGC ATCTCATTA GTCAAGCAACC ATAGTCCCGC CCCTTAATCC GCCCATCCCG
300	CCCCTAACTC CGCCCAAGTTC CGCCCCATCTT CGGCCCATATG GCTGACTTAAT TTTTTTTATT
360	TATGCAGAGG CCGAGGGCCGC CTICGGCCCTCT GAGCTATTC AGAAGTAGTG AGGAGGCTTT
420	TTTGGAGGC TAGGTTTTG CAAAAAGCTA GCTTATCCGG CGGGAAACGG TGCATTGGAA
480	CGCGGATTTC CCGTGCAG AGTGACGTAA GTACCGCTA TAGAGGACT AGTCCACCAT
540	GACCGAGTAC AAGGCCAACGG TGCGGCCCTCGC CACCCGGAC GACGTCGGCG GGGCCGTACG
600	CACCCCTCGCC GCCGCGTTCG CCGACTAACCC CGCCACGGCG CACACCGTAG ACCCGGACCG
660	CCACATCGAG CGGGTCACCG AGCTGCAAGA ACTCTTCTCT ACGGCGTCTG GGCTCGACAT
720	CGGCAAGGTG TGGGTGCGG AC6ACGGCGC CGGGGTGGCG GTCTGGACCA CGCCGGAGAG
780	CGTCGAAGCG GGGGGGGTGT TCGCCGAGAT CGGCCCGCGC ATGGCGAGT 1GAGGGGTTC
840	CGGGCTGGCC GGGCAGCAAC AGATGGAAGG CCTCCCTGGCG CGGCACCGGC CAAAGGGAGCC
900	CGCGTGGTTC CTGGCCACCG TCGGGTCTC GCCCGACAC CAGGGCAAGG GTCTGGGCAG

960 CGCCGTCGTG CTCCCCGGAG TGAGGGGG CGAGGGGG CGGGGTGCCG CCTTCCTGGGA
 1020 GACCTCCGGG CCCCCAACC TCCCCTCTA CGAGGGCTC GGCTTCACCG TCACCGCCGA
 1080 CGTCGAGTGC CGAAAGGACC GCGGACCTG GTGCATGACC CGCAAGCCG GTGCCAACAT
 1140 GGTTGACCA TTGAACTGCA TGTCGCCGT GTCCCCAAAT ATGGGGATTG GCAAGAACGG
 1200 AGACCTACCC TGCCCTCCGC TAGGAAACGC GTICAAGTAC TTCAAAAGAA TGACCAAC
 1260 CTCTTCAGTG GAAGGTAAC AGGAATCTGGT GATTATGGGT AGGGAAACCT GGTTCTCCAT
 1320 TCCTGAGAAG AATCGACCTT TAAAGGACAG AATTAAATATA GTTCTCAGTA GAGAACTCAA
 1380 AGAACCAACCA CGAGGAGCTC ATTTCCTTGC CAAAAGTTG GATGATGCCT TAAGACTTAT
 1440 TGAACAAACCG GAATTGGCAA GTAAAGTGA CATGGTTGG ATAGTGGAG GCAGTTCTGT
 1500 TTACCAAGAA GCCATGAATC AACCAAGGCCA CCTTAGACTC TTGTGACAA GGATCATGCA
 1560 GGAATTGAA AGTGCACAGT TTTCGCCAGA AATTGATTG GGGAAATAAACCTCTCCC
 1620 AGAATAACCA GGGGTCTCT CTGAGGTCCA GGAGGGAAA GGCATCAAGT ATAAGTTGA
 1680 AGTCTACGAG AAGAAAGACT AACGTTAATC GCTCCCTCC TAAAGCTATG CATTTTATA
 1740 AGACCATGGG ACTTTGCTG GCTTAGATC CCCTTGGCTT CGTGAACG CAGCTACAT
 1800 TAATACATAA CCTTATGTAT CATAACATA CGATTAGGT GACACTATAG ATAACATCCA
 1860 CTTTGCCTT CTCTCACAG GTGCCACTC CCAGGTCAA CTGCACCTCG GTTCTATCGA
 1920 TTGAATTCCA CC -Insert Sequence of Interest-
 CGATGGCC GCCATGGCCC AACTGTTA TTGCAGCTTA

Figure 14.2

TAATGGTAC AAATAAAGCA ATAGCATCAC AAATTGACA AATAAAGCAT TTTTTCACT
 GCATTCTAGT TGTGGTTGT CCAAACATCAT CAATGTACT TATCATGTCT GGATCGGGAA
 TTAATTGGC GCAGGACCAT GGCCTGAAAT AACCTCTGAA AGAGGAACCT GGTTAGGTAC
 CTTCTGAGC GGAAGAACC AGCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC
 AGGCTCCCA GCAGGCAGAA GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAAGGTG
 TGGAAAGTC CCAGGCTCCC CAGCAGGCCAG AAGTATGCCA AGCATGCATC TCAATTAGTC
 AGCAACATA GTCCCGCCCC TAATCCTGCC CATCCGCC CTAACCTCCGC CCAGGTTCCGC
 CCATTCTCCG CCCCATGGCT GACTAATTT TTTTATTAT GCAGAGGCCG AGGCCGCC
 GGCCTCTGAG CTATCCAGA AGTAGTGAGG AGGTTTTT GGAGGAGCTT TTGCAAAAG
 CTAGCTTATC CGGGGGAA CGGTGCATTG GAACGGGAT TCCCCGTGCC AAGAGTCAGG
 TAAGTACCGC CTATAGAGTC TATAGGCCCA CCCCCTTGGC TTGGTTAGAA CGGGCTACA
 ATTAATACAT AACCTTTGG ATCGATCCTA CTGACACTGA CATCCACTT TTCTTTTCT
 CCACAGGTG CCACTCCAG GTCCAACACTGC ACCTCGGTTG GCGTAGCTG CTTGGGCTGC
 ATCGATTGAA TTCCACCA -Insert Sequence of Interest-
 CGATGGCCG CATGGCCAA CTGGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT
 AGCATCACAA ATTCACAAA TAAAGCATT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC
 AAACATCA ATGTATCTTA TCAATGTCGG ATCGGGAAATT AATTGGCGC AGCACCATGG
 CCTGAAATAA GTTAAACCC TCTGAAAGAG GAAACTTGGTT AGGTACCGAC TAGTCTTTG

CAAAAGCTG TTACCTCGAG CGGCCGCTTA ATTAAGCCG GCCATTAAA TCCTGCAGGT
AACAGCTTG CACTGGCGT CGTTTACAA CGTGTGACT GGGAAACCC TGGCGTTACC
CAACTTAATC GCCTTGAGC ACATCCCCCT TTGCCAGCT GGGTAATAG CGAAGAGGCC
CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGATG GCGAATGGCG CCTGATGGG
TATTTCTTC TTACGATCT GTGCGGTATT TCACACCGCA TACGTCAAAG CAACCATAG
ACGCGCCCTG TAGCGGGCA TAAAGGGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG
CTACACTTGC CAGGGCCTA GGGCCCGCTC CTTTCGCTT CTTCCTTCC TTTCCTGCGCA
CGTTGCCGG CTTTCCCGT CAAGCTCTAA ATCGCTTGGT CCCTTTAGGG TTCCGATTAA
GTGCTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGG TGATGGTTCA CGTAGGGGC
CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGA GTCCACGTTT TTAAATAGTG
GACTCTGTT CCAAACCTGGA ACAACACTCA ACCCTATCTC GGGCTATTCT TTGATTAT
AAGGGATTG GCGATTTCG GCCTATTGGT TAAAAATGA GCTGATTAA CAAAAAATTAA
ACGGGAAATT TAACAAATA TTAACTGTTA CAATTATG GTGCACTCTC AGTACAATCT
GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT
GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCGGGGAGCT
GCATGGTCA GAGGTTTCA CGGTCACTAC CGAAACGCGC GACGAAAGGG CCTCGTGATA
CGCCTATTG TATAGGTTAA TGTCTATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT
TTTCGGGAA ATGTGGCGG AACCCCTATT TGTTTATTG TCTAAATACA TICAAATATG

TATCCGCTCA TGAGACAAATA ACCCTGATAA ATGCTCAAT AATATTGAAA AAGGAAGAGT
ATGAGTATTIC AACATTCCG TGTCGCCCTT ATTCCCTTT TTGGGGCATT TTGCCTTCTT
GTTTTGCTC ACCCAGAAC GCTGGTAAA GTAAAAGATG CTGAAGATCA GTGGGTGCA
CGAGTGGGT ACATCGAACT GGATCTCAAC AGCGGTAAAGA TCCTTGAGAG TTTTGCCTC
GAAGAACGTT TTCCAAATGAT GAGCACTTTT AAAGTCTGCA TATGTGGGC GGTATTATCC
CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTTATTCTCA GAATGACTTG
GTTGAGTACT CACCAACTCAC AGAAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGCTG CCATAAACCAT GAGTGTAAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGGTTTTTG CACAAACATGG GGGATCATGT AACTCGCCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCAAGATC
CTGTAGCAA TGGCAACAAAC GTGCGCAAAT CTATTAACCTG GCGAAACTACT TACTCTAGCT
TCCCGGCAAC AATTAAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC
TCGGCCCTTC CGGCTGGCTG GTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT
CGCGGTATCA TTGCAGGACT GGGGCCAGAT GGTAAAGCCT CCCGTATCGT AGTTATCTAC
ACGACGGGA GTCAGGAAC TATGGATGAA CGAAATAGAC AGATCGTGA GATAAGGTGCC
TCACTGATTA AGCATTTGTA ACTGTCAAGAC CAAGTTACT CATATATACT TTAGATTGAT
TTAAAACCTTC ATTTTAATT TAAAGGAGTC TAGGTGAAGA TCCTTTTGA TAATCTCATG
ACCAAAATCC CTTAACGTGA GTTTTCGTTCACTGAGCGT CAGACCCGT AGAAAAGATC

AAAGGGATCTT CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGC AAAAAAAA
CCACCGCTAC CAGCGGTGGT TTGTTTGCCT GATCAAGAGC TACCAACTCT TTTTCCGAAAG
GTAACGGCT TCAGGAGAGC GCAGATACCA AATACTGTCC TTCTAGTGT ACGGTAGTTA
GGCCACCACT TCAAGGAACCT TGTTAGCACCG CCTACATACC TCGCTCTGT AATCCTGTTA
CCAGTGGCTG CTGCCAGTGG CGATAAAGTGG TTGTCCTACCG GGTGGACTC AAGACGATAG
TTACCGGATA AGGGCAGCG GTICGGCTGA ACGGGGGTT CGTGCACACA GCCCAGCTTG
GAGCGAACGA CCTACACCGA ACTGAGATACT ACAGGGCTG AGCTATGAGA AAGCGCCACG
CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTGG AACAGGGAG
CGCACCGAGG AGCTTCCAGG GGAAAACGCC TGCTATCCTTT ATAGTCTGT CGGGTTTCGC
CACCTCTGAC TTGAGGGTGC ATTGTTGTGA TGCTCGTCAG GGGGGGGAG CCTATGGAAA
AACGCCAGCA AGGGGCCCTT TTACGGTTTC CTGGCCCTTT GCTGGCCCTT TGCTCACATG
TTCTTTCCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCCTT TGAGTGGACT
GATAACCGCTC GCGCGAGCC AACGACCGAG CGACGGAGT CAGTGGCGA GGAAGGGAA
GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCAATTATGCAGCTGG
CACGACAGGT TTCCCGACTG GAAGGGGC AGTGAGGCAC ACGCCATTAA TGTGAGTTAG
CTCACTCAT AGGCACCCCA GGCTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA
ATTGTGAGCG GATAAACATT TCAACAGGAA AACAGCTATG ACATGATTAC GAATTAA

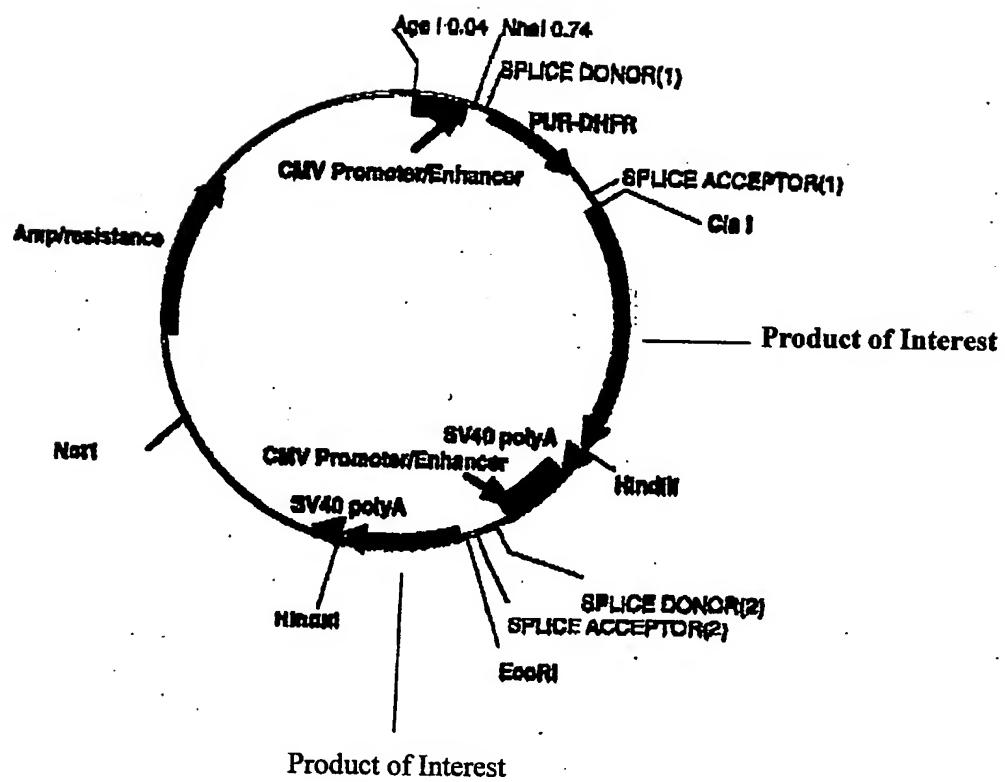


Figure 15. pCMV.IPD.HP

Timeline and Titer Comparison

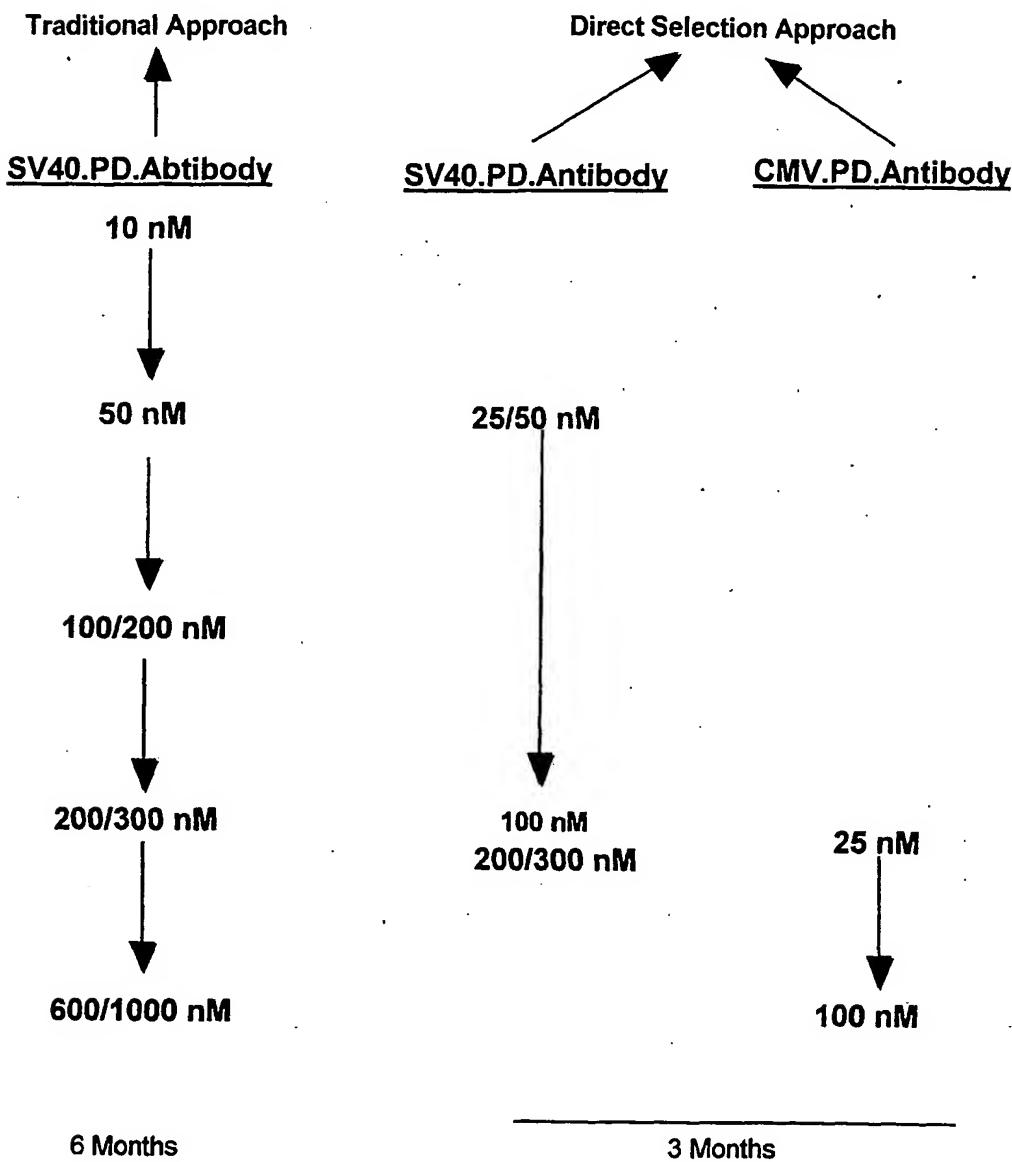


Figure 16. Timeline and Titer Comparison.

SEQUENCE LISTING

<110> Krummen, Lynne
Shen, Amy
Chisum, Venessa

<120> INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
EXPRESSING
PRODUCTION CELL LINES

<130> 22338/00101

<150> US 60/426,095
<151> 2002-11-14

<160> 4

<170> PatentIn version 3.1

<210> 1
<211> 5147
<212> DNA
<213> Artificial

<220>

<223> plasmid pSV.IPUR circular ds-DNA

<400> 1
ttcagactcg cccgacattg attattgact agagtcgatc gacagctgtg gaatgtgtgt 60
cagtttaggtt gtggaaagtc cccaggtctcc ccagcaggca gaagtatgca aagcatgcat 120
ctcaatttagt cagcaaccag gtgtggaaag tccccaggtt ccccaaggcagg cagaagtatg 180
caaagcatgc atctcaatta gtcagcaacc atagtcccgcc ccctaactcc gccccatcccg 240
ccctaactc cgcccaaggttc cgcccaattctt ccggcccatg gctgactaat tttttttattt 300
tatgcagagg ccgaggccgc ctcggctct gagctattcc agaagtatgtt aggaggcttt 360
tttggaggcc taggcttttcaaaaagcta gcttatccgg ccggaaacgg tgcattggaa 420
cgccgattcc ccgtgccaag agtgcgttaa gtaccgccta tagagcgact agtccaccat 480
gaccgagttac aagcccacgg tgcgcctcgcc cacccgcgac gacgtcccccc gggccgtacg 540
cacccctcgcc gccgcgttcg ccgactaccc cgcacgcgc cacaccgtcg accccggaccg 600
ccacatcgag cgggtcaccc agctgcaaga actcttcctc acgcgcgtcg ggctcgacat 660
cgccaagggtg tgggtcgccg acgacggcgc cgcggggcg gtctggacca cgccggagag 720
cgtcgaagcg ggggcgggtgt tcgcccagat cggccgcgc atggccgagt tgagcggttc 780
ccggctggcc ggcgcagcaac agatggaagg cctcctggcg ccgcacccggc ccaaggagcc 840
cgcggttcc ctggccacccg tcggcgtctc gcccgaccac cagggcaagg gtctggcag 900

cggcgctgt	ctccccggag	tggaggcgcc	cgagcgccc	gggggtcccc	ctttcctgga	960
gacccctcg	ccccgcaacc	tccccttcta	cgagcggtc	ggcttcaccc	tcaccgccc	1020
cgtcgagtgc	ccgaaggacc	gcccgcaccc	gtgcaccc	cgcaagcccc	gtgcctgagt	1080
taactgctcc	cctcctaaag	ctatgcattt	ttataagacc	atggacttt	tgctggcttt	1140
agatccccctt	ggcttcgtta	gaacgcagct	acaattaata	cataaccctta	tgtatcatac	1200
acatacgatt	taggtgacac	tatagataac	atccacttt	cctttctctc	cacagggtgc	1260
cactcccagg	tccaaactgca	cctcggttct	atcgattgaa	ttccccgggg	atcctctaga	1320
gtcgacctgc	agaagcttcg	atggccgcca	tggcccaact	tgtttattgc	agcttataat	1380
ggttacaaat	aaagcaatag	catcacaaat	ttcacaaata	aagcattttt	ttcactgcat	1440
tctagttgtg	gtttgtccaa	actcatcaat	gtatcttatac	atgtctggat	cgatcggaa	1500
ttaattcggc	gcagcaccat	ggcctgaaat	aacctctgaa	agaggaactt	ggttaggtac	1560
cttctgaggc	ggaaagaacc	agctgtggaa	tgtgtgtcag	ttagggtgtg	gaaagtccccc	1620
aggctccccca	gcaggcagaa	gtatgcaaag	catgcaccc	aattagtca	caaccaggta	1680
tggaaagtcc	ccaggctccc	cagcaggcag	aagtatgca	agcatgcac	tcaatttagtc	1740
agcaaccata	gtcccgcccc	taactccgccc	catccgccc	ctaactccgc	ccagttccgc	1800
ccattctccg	ccccatggct	gactaatttt	ttttatattat	gcagaggccg	aggccgcctc	1860
ggcctctgag	ctattccaga	agtagtgagg	aggctttttt	ggaggcctag	gcttttgcaa	1920
aaagctgtta	cctcgagcgg	ccgcttaatt	aaggcgccc	attnaatcc	tgcaggtaac	1980
agcttggcac	tggccgtcg	tttacaacgt	cgtgactgg	aaaaccctgg	cgttacccaa	2040
cttaatcgcc	ttgcagcaca	tcccccttc	gccagctggc	gtaatagcga	agaggcccgc	2100
accgatcgcc	cttcccaaca	gttgcgtagc	ctgaatggcg	aatggccct	gatgcggat	2160
tttctcccta	cgcacatgtg	cggatattca	caccgcatac	gtcaaagcaa	ccatagtagc	2220
cgcctgttag	cggcgcatta	agcgccgg	gtgtgggtgt	tacgcgcagc	gtgaccgcta	2280
cacttgcag	cgccttagcg	ccgcgtcctt	tgcgtttttt	cccttcctt	ctgcacgt	2340
tgcggcgtt	tccccgtcaa	gctctaaatc	gggggtcccc	tttagggttc	cgattttagtg	2400
cttacggca	cctcgacccc	aaaaaacttg	atttgggtga	tggttcacgt	agtggccat	2460
cgcctgtata	gacggttttt	cgcctttga	cggtggagtc	cacgttcttt	aatagtggac	2520
tcttgttcca	aactggaaca	acactcaacc	ctatctcg	tttattttt	gatttataag	2580

ggattttgcc gatttcggcc tattggtaa aaaatgagct gatthaacaa aaatttaacg	2640
cgaattttaa caaaatatta acgttacaa tttatggtg cactctcagt acaatctgct	2700
ctgatgccgc atagttaaac caactccgct atcgctacgt gactgggtca tggctgcgc	2760
ccgacaccccg ccaacaccccg ctgacgcgc ctgacgggct tgtctgctcc cggcatccgc	2820
ttacagacaa gctgtgaccg tctccggag ctgcatgtgt cagaggtttt caccgtcatac	2880
accgaaacgc gcgaggcagt attcttgaag acgaaaggc ctcgtatac gcctat	2940
ataggtaat gtcatgataa taatggttt ttagacgtca ggtggcactt ttcgggaaa	3000
tgtgcgcgga acccctat ttttattttt ctaaatacat tcaaataatgt atccgctcat	3060
gagacaataa ccctgataaa tgctcaata atattgaaaa aggaagagta tgagtattca	3120
acatcccgt gtcgcctta ttccctttt tgcggcattt tgcctcctg ttttgc	3180
cccgagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta	3240
catcgactg gatctcaaca gcggtaagat ctttgcgc aagaacgttt	3300
tccaatgatg agcactttt aagttctgct atgtggcgcgtt gtattatccc gtgatgacgc	3360
cggcaagag caactcggcgc gcccataca ctattctcag aatgacttgg ttgagtactc	3420
accagtcaca gaaaagcatac ttacggatgg catgacagta agagaattat gcagtgc	3480
cataaccatg agtgataaca ctgcggccaa ctacttctg acaacgatcg gaggaccgaa	3540
ggagctaacc gctttttgc acaacatggg ggatcatgta actcgcccttgc atcggtgggaa	3600
accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc cagcagcaat	3660
ggcaacaacg ttgcgcacac tattaactgg cgaactactt actctagctt cccggcaaca	3720
attaatagac tggatggagg cggataaaatg tgcaggacca cttctgcgtc cggcccttcc	3780
ggctggctgg ttatttgcgataa ataaatctgg agccggtagt cgtgggtctc gcggtatcat	3840
tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag	3900
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa	3960
gcattggtaa ctgtcagacc aagtttactc atataactt tagattgatt taaaacttca	4020
tttttaat taaaaggatct aggtgaagat ctttttgcgat aatctcatga ccaaaaatccc	4080
ttaacgtgag ttttgcgtcc actgagcgtc agaccccgta gaaaagatca aaggatctt	4140
ttgagatcct tttttctgc gcgtaatctg ctgcttgcgaa aaaaaaaaaac caccgctacc	4200
agcgggtggtt tttttgcggg atcaagagct accaactt tttccgaagg taactggctt	4260
cagcagagcg cagataccaa atactgtcct tcttagtgcgtt ccgtatgttgc accaccactt	4320

caagaactct	gtagcaccgc	ctacataacct	cgctctgcta	atccctgttac	cagtggctgc	4380
tgccagtgcc	gataagtcgt	gtcttaccgg	gttggactca	agacgatagt	taccggataa	4440
ggcgcagcgg	tcgggctgaa	cgggggggttc	gtgcacacag	cccagctgg	agcgaacgac	4500
ctacaccgaa	ctgagatacc	tacagcgtga	gcattgagaa	agcgccacgc	ttcccgaagg	4560
gagaaaggcg	gacaggtatc	cggtaagcgg	cagggtcgga	acaggagagc	gcacgaggga	4620
gcttccaggg	ggaaacgcct	ggtatcttta	tagtcctgtc	gggttcgcc	acctctgact	4680
tgagcgtcga	tttttgcgtat	gctcggtcagg	ggggcggagc	ctatggaaaa	acgcccagcaa	4740
cgcggcctt	ttacggttcc	tggcccttttgc	ctggccttttgc	gctcacatgt	tcttcctgc	4800
gttatcccct	gattctgtgg	ataaccgtat	taccgcctt	gagtgagctg	ataccgctcg	4860
ccgcagccga	acgaccgagc	gcagcgagtc	agtgagcggag	gaagcggaaag	agcgc当地	4920
acgcaaaccg	cctctccccc	cgcgttggcc	gattcattaa	tccagctggc	acgacaggtt	4980
tcccgactgg	aaagcgggca	gtgagcgcaa	cgcaattaat	gtgagttacc	tcactcatta	5040
ggcacccca	gctttacact	ttatgttcc	ggctcgtatg	ttgtgtggaa	ttgtgagcgg	5100
tattgttaaa	gtgtgtcctt	tgtcgatact	ggtactaatg	cttaatt		5147

<210> 2
<211> 5171
<212> DNA
<213> Artificial

<220>
<223> plasmid pSV.ID circular ds-DNA

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

<220>
<221> misc_feature
<222> (1946)..(1946)
<223> start pUC118

<220>
<221> misc_feature
<222> (1954)..(1954)
<223> linearization linker inserted into Hpal site

<220>

<221> misc_feature
 <222> (529)..(1090)
 <223> DHFR coding region

<220>
 <221> misc_feature
 <222> (1522)..(1522)
 <223> sv40 origin

<400> 2						
ttcgagctcg	cccgacattg	attattgact	agagtcgatc	gacagctgtg	aatgtgtgt	60
cagtttagggt	gtggaaagtc	cccaggctcc	ccagcaggca	gaagtatgca	aagcatgcat	120
ctcaattagt	cagcaaccag	gtgtggaaag	tccccaggct	ccccagcagg	cagaagtatg	180
caaagcatgc	atctcaatta	gtcagcaacc	atagtccgc	ccctaactcc	gcccatcccg	240
ccccctaactc	cggccaggttc	cggccattct	ccggcccatg	gctgactaat	tttttttatt	300
tatgcagagg	ccgaggccgc	ctcggcctct	gagctattcc	agaagtagtg	aggaggctt	360
tttggaggcc	taggctttt	caaaaagcta	gcttatccgg	ccgggaacgg	tgcattggaa	420
cgcggattcc	ccgtgccaag	agtgacgtaa	gtaccgccta	tagatctat	aggcccaccc	480
cttggctcta	gagagatata	agcctaggat	tttatcccg	gtgccatcat	ggttcgacca	540
ttgaactgca	tcgtcgccgt	gtcccaaata	atggggattt	gcaagaacgg	agacctaccc	600
tgcctccgc	tcaaggaacgc	gttcaagtac	ttccaaagaa	tgaccacaac	ctttcagtg	660
gaaggtaaac	agaatcttgt	gattatgggt	aggaaaacct	ggttctccat	tcctgagaag	720
aatcgacctt	taaaggacag	aattaatata	gttctcagta	gagaactcaa	agaaccacca	780
cgaggagctc	atttcttgc	caaaagtttt	gatgatgcct	taagacttat	tgaacaaccc	840
gaattggcaa	gtaaagtaga	catggtttgg	atagtcggag	gcagttctgt	ttaccaggaa	900
gccatgaatc	aaccaggcca	ccttagactc	tttgtgacaa	ggatcatgca	ggaatttcaa	960
agtgacacgt	ttttcccaga	aattgatttg	ggaaatata	aacctctccc	agaataccca	1020
ggcgtcctct	ctgaggtcca	ggaggaaaaaa	ggcatcaagt	ataagttga	agtctacgag	1080
aagaaagact	aacaggaaga	tgctttcaag	ttctctgctc	ccctccctaa	gctatgcatt	1140
tttataagac	catgggactt	ttgctggctt	tagacccct	tggcttcgtt	agaacgcggc	1200
tacaattaat	acataaccctt	atgtatcata	cacatagatt	tagtgacac	tatagaataaa	1260
catccacttt	gcctttctct	ccacaggtgt	cactccagg	caactgcacc	tccggttctat	1320
cgattgaatt	ccccggggat	cctctagagt	cgacctgcag	aagcttggcc	gccccatggccc	1380

aacttgttta ttgcagctta taatggttac aaataaaagca atagcatcac aaatttcaca	1440
aataaagcat tttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct	1500
tatcatgtct ggatcgatcg ggaattaatt cggcgagca ccatggcctg aaataaacctc	1560
tgaaaagagga acttggtag gtaccttctg aggcgaaag aaccagctgt ggaatgtgtg	1620
tcaatgggg tgtggaaagt ccccaggctc cccagcagc agaagtatgc aaagcatgca	1680
tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag gcagaagttat	1740
gcaaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cggccatccc	1800
ccccctaact cggcccatgtt ccgccccatc tccgccccat ggctgactaa ttttttttat	1860
ttatgcagag gccgaggccg cctccggctc tgagctattc cagaagtagt gaggaggctt	1920
ttttggagc ctaggctttt gcaaaaagct gttacctcga gcggccgctt aattaaggcg	1980
cgccatttaa atccctgcagg taacagcttg gcactggccg tcgttttaca acgtcgtgac	2040
tggaaaaacc ctggcggttac ccaacttaat cgccttgcag cacatcccc cttcgccagc	2100
tggcgtaata gcgaagaggc ccgcaccgat cgccttcccc aacagttgcg tagcctgaat	2160
ggcgaatggc gcctgatgcg gtattttctc cttacgcattc tgcgtttat ttcacaccgc	2220
atacgtcaaa gcaaccatag tacgcgcctt gtgcggcgc attaagcgcg gcgggtgtgg	2280
tggttacgcg cagcgtgacc gctacacttg ccagcgcctt agcgcgcgtt ctttcgcatt	2340
tcttcccttc ctttctcgcc acgttcgcgg gctttcccc tcaagctcta aatcgggggc	2400
tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgatttgg	2460
gtgatggttc acgttagtggg ccacgcgcctt gatagacggt ttttgcgcct ttgacgttgg	2520
agtccacgtt cttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct	2580
cgggcatttc ttttgcattta taaggattt tgccgatttc ggcctattgg ttaaaaaatg	2640
agctgattta acaaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaattttat	2700
ggtcactct cagtacaatc tgctctgatc ccgcatacggtt aagccaaactc cgctatcgct	2760
acgtgactgg gtcatggctg cggccgcaca cccgccaaca cccgctgacg cggccatgcg	2820
ggcttgcgtg ctcccgcat ccgcattacag acaagctgtg accgtctccg ggagctgcatt	2880
gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg cagtattttt gaagacgaaa	2940
gggcctcggtt atacgcctat ttttatacggtt taatgtcatg ataataatgg tttcttagac	3000
gtcaggtggc acttttcggg gaaatgtgcg cggaaaccctt atttgtttat tttctaaat	3060

acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg	3120
aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct ttttgcggc	3180
atttgcctt cctgttttg ctcacccaga aacgctggg aaagtaaaag atgctgaaga	3240
tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga	3300
gagtttcgc cccgaagaac gtttccaat gatgagcaact tttaaagttc tgctatgtgg	3360
cgccgttatta tcccgtgatg acgcccggca agagcaactc ggtcgccgca tacactattc	3420
tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac	3480
agtaagagaa ttatgcagtg ctgccataac catgagtgt aacactgcgg ccaacttact	3540
tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca	3600
tgtaactcgc cttgatcggtt gggAACCGGA gctgaatgaa gccataccaa acgacgagcg	3660
tgacaccacg atgccagcag caatggcaac aacgttgcgc aaactattaa ctggcgaact	3720
acttactcta gcttccccc aacaattaat agactggatg gaggcggata aagttgcagg	3780
accacttctg cgctcgcccc ttccggctgg ctggtttatt gctgataaat ctggagccgg	3840
tgagcgtggg tctcgcgta tcattgcagc actggggcca gatggtaagc cctcccgat	3900
cgtagttatc tacacgacgg ggagtcaggg aactatggat gaacgaaata gacagatcgc	3960
tgagataggt gcctcactga ttaagcattt gtaactgtca gaccaagttt actcatatat	4020
acttttagatt gatttaaaac ttcatttta atttaaaagg atctaggtga agatccttt	4080
tgataatctc atgacaaaaa tcccttaacg ttagtttgc ttccactgag cgtagacccc	4140
cgtagaaaag atcaaaggat cttcttgaga tcctttttt ctgcgcgtaa tctgctgctt	4200
gcaaacaaaa aaaccaccgc taccagcggt gtttttttgc ccggatcaag agctaccaac	4260
tcttttccg aaggttaactg gtttcagcag agcgcagata ccaaatactg tccttctagt	4320
gtagccgtag tttaggcacc acttcaagaa ctctgttagca ccgcctacat acctcgctct	4380
gctaattcctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgg	4440
ctcaagacga tagttacccg ataaggcgca gcgggtcgccc tgaacggggg gttcgtgcac	4500
acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagcattt	4560
agaaaagcgcc acgcttcccg aaggagaaa ggcggacagg tatccgtaa gcggcagggt	4620
cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatac tttatagtcc	4680
tgtcgggttt cggccacctct gacttgagcg tcgatttttt tgatgctgtt cagggggcg	4740
gaggctatgg aaaaacgcca gcaacgcggc ctttttacgg ttccctggcct tttgctggcc	4800

tttgctcac atgttcttc ctgcgttata ccctgattct gtggataacc gtattaccgc 4860
cttgagtga gctgataaccg ctcgccgcag ccgaacgacc gagcgcagcg agtcagttag 4920
cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt gcccgattca 4980
ttaatccagc tggcacgaca ggttcccga ctggaaagcg ggcagtgagc gcaacgcaat 5040
taatgtgagt tacctcactc attaggcacc ccaggctta cacttatgc ttccggctcg 5100
tatgttgtgt ggaattgtga gcggataaca atttcacacaca ggaaacagct atgaccatga 5160
ttacgaatta a 5171

<210> 3
<211> 5712
<212> DNA
<213> Artificial

<220>
<223> plasmid pSV.IPD circular ds-DNA

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

<220>
<221> misc_feature
<222> (479)..(479)
<223> start PUR coding

<220>
<221> misc_feature
<222> (1079)..(1643)
<223> DHFR coding region

<400> 3
ttcgagctcg cccgacattg attattgact agagtcgatc gacagctgtg gaatgtgtgt 60
cagttagggt gtggaaagtc cccagggctcc ccagcaggca gaagtatgca aagcatgcat 120
ctcaattagt cagcaaccag gtgtggaaag tccccaggt ccccaaggcagg cagaagtatg 180
caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccattccg 240
cccctaactc cgcccagttc cgccccattct ccgccccatg gctgactaat tttttttatt 300
tatgcagagg ccgaggccgc ctggccctct gagctattcc agaagtagtg aggaggctt 360
tttggaggcc taggcttttg caaaaagcta gcttatccgg ccggaaacgg tgcattggaa 420

cgccgattcc ccgtgccaag agtgcgttaa gtaccgccta tagagcgact agtccaccat	480
gaccgagtac aagcccacgg tgccgcctcg caccgcgcac gacgtcccgc gggccgtacg	540
caccctcgcc gccgcgttgc ccgactaccc cgccacgcgc cacaccgtag accccggaccg	600
ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat	660
cggcaagggtg tgggtcgccg acgacggcgc cgcgggtggcg gtctggacca cgccggagag	720
cgtcgaagcg ggggcgggtgt tcgcccagat cggccgcgc atggccgagt tgagcggttc	780
ccggctggcc gcgcagcaac agatggaagg ctcctggcg ccgcaccggc ccaaggagcc	840
cgcgtggttc ctggccaccg tcggcgttc gcccgaccac cagggcaagg gtctggcag	900
cgcgtcggtg ctccccggag tggaggcggc cgagcgcgc ggggtgccccg cttcctgga	960
gacctccgcg ccccgcaacc tcccttcta cgagcggctc ggcttcaccg tcaccgcga	1020
cgtcgagtgc ccgaaggacc gcgcgacctg gtgcattgacc cgcaagccccgtgccaacat	1080
ggttcgacca ttgaactgca tcgtcgccgt gtcccaaataat atggggattt gcaagaacgg	1140
agacctaccc tgccctccgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac	1200
ctcttcagtg gaaggtaaac agaatctggt gattatgggt aggaaaacctt gttctccat	1260
tcctgagaag aatcgaccctt taaaggacag aattaatata gttctcagta gagaactcaa	1320
agaaccacca cgaggagctc atttcttgc caaaagtttgc gatgtgcct taagacttat	1380
tgaacaaccg gaatttggcaa gtaaagttaga catggtttgg atagtcggag gcagttctgt	1440
ttaccaggaa gccatgaatc aaccaggcca ctttagactc ttttgacaa ggatcatgca	1500
ggaatttggaa agtgcacacgt ttttcccaga aattgatttgc gggaaatata aacctctccc	1560
agaataacca ggcgtccctct ctgaggtcca ggaggaaaaa ggcataagt ataagtttga	1620
agtctacgag aagaaagact aacgttaact gctccctcc taaagctatg cattttata	1680
agaccatggg acttttgcgt gcttttagatc cccttggctt cgttagaacg cagctacaat	1740
taatacataa ctttatgtat catacacata cgattnaggat gacactatag ataacatcca	1800
ctttgccttt ctctccacag gtgtccactc ccaggtccaa ctgcacccctcg gttctatgca	1860
ttgaattccc cggggatcct ctagagtcga cctgcagaag cttcgatggc cgccatggcc	1920
caacttggttt attgcagttt ataatggta caaataaagc aatagcatca caaatttcac	1980
aaataaagca ttttttcac tgcattctag ttgtggtttgc tccaaactca tcaatgtatc	2040
ttatcatgtc tggatcgatc ggaattaaat tcggcgcagc accatggcct gaaataacct	2100
ctgaaagagg aacttggtta ggtaccttctt gaggcggaaa gaaccagctg tggaaatgtgt	2160

gtcagttagg gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc	2220
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta	2280
tgcaaagcat gcatctcaat tagtcagcaa ccatagtcgg ccccttaact cgcggccatcc	2340
cgcggccatcc tccggccagg tccggccatt ctccggccca tggctgacta atttttttta	2400
tttatgcaga ggccgaggcc gcctcgccct ctgagctatt ccagaagtag tgaggaggct	2460
ttttggagg cctaggctt tgcaaaaagc tgttacctcg agcggccgct taattaaggc	2520
gcgcattta aatcctgcag gtaacagctt ggcactggcc gtcgttttac aacgtcgtga	2580
ctggggaaaac cctggcgta cccaaattaa tcgccttgca gcacatcccc cttcgccag	2640
ctggcgtaat agcgaagagg cccgcaccga tcgccttcc caacagttgc gtagcctgaa	2700
tggcgaatgg cgccctgatgc ggtatttct ccttacgcat ctgtgcggta tttcacaccg	2760
catacgtcaa agcaaccata gtacgcgccc ttagcggcg cattaagcgc ggcgggtgt	2820
gtggttacgc gcagcgtgac cgctacactt gccagcgcgg tagcgcggc tccttcgct	2880
ttcttcctt ctttctcgcc cacgttcgccc ggcttcccc gtcaagctct aaatcggggg	2940
ctccctttag ggttccgatt tagtgcattt cggcacctcg accccaaaaa acttgatttg	3000
ggtgatggtt cacgttagtgg gccatcgccc ttagatagacgg ttttcgccc tttgacgttg	3060
gagttccacgt tcttaatag tggactcttgc ttccaaactg gaacaacact caaccctatc	3120
tcgggctatt ctittgattt ataaggatt ttgccgattt cggcctattt gttaaaaat	3180
gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgtt tacaattttt	3240
tggtgactc tcagtacaat ctgctctgat gccgcatagt taagccact ccgtatcg	3300
tacgtgactg ggtcatggct gcgcggcggac acccgccaaac acccgctgac gcgcctgac	3360
gggcttgcgt gctccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca	3420
tgtgtcagag gtttccaccc tcacaccga aacgcgcgag gcagtattct tgaagacgaa	3480
agggcctcggt gatacgccta ttttatagg ttaatgtcat gataataatg gtttctttaga	3540
cgtcagggtgg cacttttcgg ggaaatgtgc gcggaaacccc tatttgcgtt ttttctaa	3600
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt	3660
aaaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc tttttgcgg	3720
cattttgcct tcctgtttt gtcacccag aaacgctggt gaaagtaaaa gatgctgaag	3780
atcagttggg tgcacgagtg gtttacatcg aactggatct caacagcggt aagatccttgc	3840

agagtttcg ccccgaagaa cgtttccaa ttagtggcac tttaaagtt ctgttatgtg	3900
gcgcggatt atcccgat gacccgggc aagagcaact cggcgccgc atacactatt	3960
ctcagaatga ctgggttag tactcaccag tcacagaaaa gcatcttacg gatggcatga	4020
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac	4080
ttctgacaac gatcggagga ccgaaggagc taaccgctt ttgcacaac atgggggatc	4140
atgttaactcg ctttgcgtg tggaaaccgg agctgaatga agccatacca aacgacgagc	4200
gtgacaccac gatgccagca gcaatggcaa caacgttgcg caaactatta actggcgaac	4260
tacttactct agttcccg caacaattaa tagactggat ggaggcggat aaagttgcag	4320
gaccacttct gcgcctggcc cttccggctg gctggttat tgctgataaa tctggagccg	4380
gtgagcgtgg gtctcgccgt atcattgcag cactggggcc agatggtaag ccctcccgta	4440
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg	4500
ctgagatagg tgcctcactg attaaggcatt ggttaactgtc agaccaagtt tactcatata	4560
tacttttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccctt	4620
ttgataatct catgaccaaa atcccttaac gtgagtttc gttccactga gcgtcagacc	4680
ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct	4740
tgcaaacaaa aaaaccaccc ctaccagccg tggttgtt gccggatcaa gagctaccaa	4800
ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag	4860
tgtagccgt a ttaggccac cactcaaga actctgtac accgcctaca tacctcgctc	4920
tgctaattct gttaccagtg gctgctgca gtggcgataa gtcgtgtctt accgggttgg	4980
actcaagacg atagttaccc gataaggcgc acgcgtcggg ctgaacgggg gttcgtgca	5040
cacagccccag ctggagcga acgacctaca ccgaactgag atacctacag cgtgagcatt	5100
gagaaagcgc cacgcttccc gaagggagaa aggccggacag gatccggta agccgcagg	5160
tcggaacagg agagcgcacg agggagctc cagggggaaa cgcctggat ctttatagtc	5220
ctgtcgggtt tcgcccaccc tgacttgagc gtcgatttt gtgtatgtcg tcaggggggc	5280
ggagcctatg gaaaaacgcc agcaacgcgg ctttttacg gttccctggcc ttttgcggc	5340
ctttgctca catgttctt cctgcgttat cccctgattc tgtggataac cgtattaccc	5400
cctttgagtg agctgataacc gctcgccgca gccgaacgcac cgagcgcagc gagtcagtga	5460
gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgccggt tggccgattc	5520
attaatccag ctggcacgac aggttcccg actggaaagc gggcagttag cgcacgcacaa	5580

ttaatgtgag ttacctcaact cattaggcac cccaggcttt acactttatg ctccggctc 5640
gtatgttgta tggaatttgta agcgataac aatttcacac agaaacagc tatgaccatg 5700
attacgaatt aa 5712

<210> 4
<211> 12514
<212> DNA
<213> Artificial

<220>
<223> plasmid pSV.IPD.2C4 circular ds-DNA

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

<220>
<221> misc_feature
<222> (479)..(479)
<223> start PUR coding

<220>
<221> misc_feature
<222> (1079)..(1643)
<223> DHFR coding region

<220>
<221> misc_feature
<222> (1883)..(1883)
<223> start 2C4 HC coding

<220>
<221> misc_feature
<222> (4154)..(4154)
<223> start LC coding

<400> 4
ttcgagctcg cccgacattg attattgact agagtcgatc gacagctgtg gaatgtgtgt 60
cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
ctcaatttagt cagcaaccag gtgtggaaag tccccaggct ccccagcagg cagaagtatg 180
caaagcatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccatcccg 240
cccttaactc cgcccagttc cgcccatctt ccgccccatg gctgactaat tttttttatt 300

tatgcagagg ccgaggccgc ctcggcctct gagctattcc agaagtagtg aggaggctt	360
tttggaggcc taggcttttgc caaaaagcta gcttatccgg ccgggaacgg tgcattggaa	420
cgccgattcc ccgtgccaag agtgacgtaa gtaccgccta tagagcgact agtccaccat	480
gaccgagttac aagcccacgg tgcgcctcgc caccgcgcac gacgtcccgac gggccgtacg	540
caccctcgcc gccgcgttcg ccgactaccc cgccacgcgc cacaccgtag acccggaccg	600
ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat	660
cggcaagggtg tgggtcgcgg acgacggcgc cgccgtggcg gtctggacca cgccggagag	720
cgtcgaagcg gggcggtgt tcgcccagat cggcccgccg atggccgagt tgagcggttc	780
ccggctggcc ggcgcagcaac agatggaagg ctcctggcg ccgcaccggc ccaaggagcc	840
cgcgtggttc ctggccaccg tcggcgtctc gcccgcaccac cagggcaagg gtctggcag	900
cgccgtcggtg ctccccggag tggaggcggc cgagcgcgcg ggggtgccccg cttccctgg	960
gacctccgcg ccccgcaacc tcccctcta cgagcggctc ggcttcaccg tcaccgcga	1020
cgtcgtgcgcc ccccgcaacc tcccctcta cgagcggctc ggcttcaccg tcaccgcga	1080
ggttcgacca ttgaactgca tcgtcgccgt gtcccaaataat atggggatttgc gcaagaacgg	1140
agacctaccc tggccctccgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac	1200
ctcttcagtg gaaggtaaac agaatctggt gattatgggt aggaaaacct ggttctccat	1260
tcctgagaag aatcgaccc taaaggacag aattaatata gttctcagta gagaactcaa	1320
agaaccacca cgaggagctc attttcttgc caaaaagtttgc gatgtgcct taagacttat	1380
tgaacaacccg gaattggcaa gtaaaagttaga catggtttgg atagtcggag gcagttctgt	1440
ttaccaggaa gccatgaatc aaccaggcca ctttagactc tttgtgacaa ggatcatgca	1500
ggaatttgaa agtacacgt ttttccaga aattgatttg gggaaatata aacctctccc	1560
agaataacca ggcgtcctct ctgaggtcca ggaggaaaaa ggcataagt ataagtttg	1620
agtctacgag aagaaagact aacgttaact gctccctcc taaagctatg cattttata	1680
agaccatggg acttttgcgt gtttagatc cccttggctt cgtagaaacg cagctacaat	1740
taatacataa ctttatgtat catacacata cgattttaggt gacactatag aataacatcc	1800
actttgcctt tctctccaca ggtgtccact cccaggtcca actgcacccctc ggttctatcg	1860
attgaattcc accatgggat ggtcatgtat catcctttt ctagtagcaa ctgcaactgg	1920
agtacattca gaagttcagc tgggtggagtc tggcggtggc ctgggtgcagc cagggggctc	1980

actccgttg tcctgtgcag cttctggctt cacccacc gactatacca tggactgggt	2040
ccgtcaggcc ccgggtaagg gcctggaatg gttgcagat gttatcccta acagtggcgg	2100
ctctatctat aaccagcgct tcaagggccg tttcactctg agtggata gatctaaaaa	2160
cacattatac ctgcagatga acagcctgcg tgctgaggac actgccgtct attattgtgc	2220
tcgtaacctg ggacccttctt tctactttga ctactgggtt caaggaaccc tggtcaccgt	2280
ctccctggcc tccaccaagg gcccatcggt cttcccccctg gcaccctcct ccaagagcac	2340
ctctggggcc acagcggccc tgggtgcct ggtcaaggac tacttcccg aaccggtgac	2400
gggtgcgtgg aactcaggcg ccctgaccag cggcgtgcac accttcccg ctgtcctaca	2460
gtcctcagga ctctactccc tcagcagcgt ggtgactgtg ccctctagca gcttgggcac	2520
ccagacctac atctgcaacg tgaatcacaa gcccagcaac accaaggtgg acaagaaaagt	2580
tgagccaaa tcttgtgaca aaactcacac atgcccaccc tgcccagcac ctgaactcct	2640
ggggggaccc tcagtcttcc tcttccccc aaaacccaag gacaccctca tgcatacccg	2700
gaccctgag gtcacatgctg tgggtgtgg cgtgagccac gaagaccctg aggtcaagtt	2760
caactggta cgtggacggcg tggaggtgca taatgccaag acaaagccgc gggaggagca	2820
gtacaacagc acgtaccggg tggtcagcgt cctcaccgtc ctgcaccagg actggctgaa	2880
tggcaaggag tacaagtgca aggtctccaa caaagccctc ccagccccc tgcagaaaaac	2940
catctccaaa gccaaaggc agccccgaga accacaggtg tacaccctgc ccccatcccg	3000
ggaagagatg accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccg	3060
cgacatcgcc gtggagtggg agagcaatgg gcagccggag aacaactaca agaccacgac	3120
tcccgtgtc gactccgacg gtccttctt cctctacagc aagctcaccg tggacaagag	3180
caggtggcag caggggaacg tcttctcatg ctccgtgtatg catgaggctc tgcacaacca	3240
ctacacgcag aagagcctct ccctgtctcc gggtaatga gtgcacggc cctagagtcg	3300
acctgcagaa gcttcgtatgg ccgcacatggc ccaacttggt tattgcagct tataatggtt	3360
acaataaaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta	3420
gttgggttt gtccaaactc atcaatgtat cttatcatgt ctggatcggg aattaattcg	3480
gcgcagcacc atggcctgaa ataacctctg aaagaggaac ttggtaggt accttctgag	3540
gcggaaaagaa ccagctgtgg aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc	3600
cagcaggcag aagtatgcaa agcatgcac tcaatttagtc agcaaccagg tgtggaaagt	3660
ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca	3720

tagtcccccc	ccttaactccg	cccatccgc	ccctaactcc	gcccagttcc	gcccattctc	3780
cgccccatgg	ctgactaatt	tttttattt	atgcagaggc	cgaggccgccc	tcggcctctg	3840
agctattcca	gaagtagtga	ggaggctttt	ttggaggact	aggctttgc	aaaaagctag	3900
cttatccggc	cgggaacggt	gcattggAAC	gcccattccc	cgtgccaaga	gtcaggtaag	3960
taccgcctat	agagtctata	ggcccacccc	cttggcttcg	ttagaacgcg	gctacaatta	4020
atacataacc	ttttggatcg	atcctactga	cactgacatc	cacttttct	ttttctccac	4080
aggtgtccac	tcccaggtcc	aactgcacct	cggttcgcga	agctagttg	ggctgcatcg	4140
attgaattcc	accatggat	ggtcatgtat	catccttttt	ctagtagcaa	ctgcaactgg	4200
agtacattca	gatatccaga	tgacccagtc	cccagctcc	ctgtccgcct	ctgtggcga	4260
tagggtcacc	atcacctgca	aggccagtc	ggatgtgtct	attgggtcgc	cctggtatca	4320
acagaaacca	ggaaaagctc	cgaaactact	gatttactcg	gcttcctacc	gatacactgg	4380
agtcccttct	cgcttctctg	gatccggttc	tgggacggat	ttcactctga	ccatcagcag	4440
tctgcagcca	gaagacttcg	caacttatta	ctgtcaacaa	tattatattt	atccttacac	4500
gtttggacag	ggtaccaagg	tggagatcaa	acgaactgtg	gctgcaccat	ctgtcttcat	4560
cttcccgcca	tctgatgagc	agttgaatc	tggaactgct	tctgttgtgt	gcctgctgaa	4620
taacttctat	cccagagagg	ccaaagtaca	gtggaggtg	gataacgccc	tccaatcggg	4680
taactcccag	gagagtgtca	cagagcagga	cagcaaggac	agcacctaca	gcctcagcag	4740
caccctgacg	ctgagcaaag	cagactacga	gaaacacaaa	gtctacgcct	gccaagtcac	4800
ccatcaggc	ctgagctcgc	ccgtcacaaa	gagttcaac	aggggagagt	gttaagcttc	4860
gatggccgccc	atggcccaac	ttgtttattt	cagttataaa	tggttacaaa	taaagcaata	4920
gcatcacaaa	tttcacaaaat	aaagcatttt	tttcaactgca	ttctagttgt	ggtttgcaca	4980
aactcatcaa	tgtatcttat	catgtctgga	tcgggaatta	attcggcgca	gcaccatggc	5040
ctgaaataag	tttaaaccct	ctgaaagagg	aacttggta	ggtaccgact	agtagcaagg	5100
tcgcccacgca	caagatcaat	attaacaatc	agtcatctct	ctttagcaat	aaaaaggtga	5160
aaaattacat	tttaaaaatg	acaccataga	cgatgtatga	aaataatcta	cttggaaata	5220
aatctaggca	aagaagtgca	agactgttac	ccagaaaact	tacaaattgt	aaatgagagg	5280
ttagtgaaga	tttaaatgaa	tgaagatcta	aataaactta	taaattgtga	gagaaattaa	5340
tgaatgtcta	agttaatgca	gaaacggaga	gacatactat	attcatgaac	taaaagactt	5400

aatatttgta aggtatactt tcttttcaca taaatttgta gtcaatatgt tcaccccaa	5460
aaagctgttt gttaacttgt caacctcatt tcaaaatgtat tataaaaaagc ccaaagacaa	5520
taacaaaaat attcttgttag aacaaaatgg gaaagaatgt tccactaaat atcaagattt	5580
agagcaaaagc atgagatgtg tggggataga cagtgggct gataaaatag agtagagctc	5640
agaaaacagac ccattgat atgtaaatgtc cctatgaaaa aaatatggca ttttacaatg	5700
ggaaaatgtat gatcttttc ttttttagaa aaacaggaa atatatttat atgtaaaaaa	5760
taaaaggaa cccatatgtc ataccatatac cacaaaaaaa ttccagtgaa ttataagtct	5820
aaatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat gccatcatga	5880
cttcagtgtt gaaaaattt tcttatgact caaagtccta accacaaaga aaagattgtt	5940
aatttagattt catgaatattt aagacttattt tttaaaattt aaaaaccattt aagaaaagtc	6000
aggccataga atgacagaaa atatggcaa caccggatgt aagagaattt gtaatatgcag	6060
attataaaaaa gaagtcttac aaatcgat taaaataaaac tagacaaaaaa tttgaacaga	6120
tgaaagagaa actctaaata atcattacac atgagaaact caatctcaga aatcagagaa	6180
ctatcattgc atatacacta aatttagagaa atattaaaag gctaagtaac atctgtggca	6240
atattgtgg tatataacct tggatgtatg tggatgagaac agtacttttac cccatggct	6300
tcctcccaa acccttaccc cagttataat catgacaaat atacttttac aaccattacc	6360
ctatatctaa ccagttactcc tcaaaactgtt caaggtcatc aaaaataaga aaagtctgag	6420
gaactgtcaa aactaagagg aacccaaagg gacatgagaa ttatgtttaa tgtggcattc	6480
tgaatgagat cccagaacag aaaaagaaca gtagctaaa aactaatgaa atataaataa	6540
agtttgaact ttagttttt tttttttttt gtagcattaa cacggcaaaag tcattttcat	6600
atttttctt aacattaatgtt acaagtctat aattttttttt tttttttttt tagtctggaa	6660
cattgccaga aacagaagta cagcagctat ctgtgtgtc gccttaactat ccatacgatgt	6720
ttggtctaaa atgagataca tcaacgctcc tccatgtttt ttgtttttttt tttttttttt	6780
aaactttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	6840
aagctgagga agttttccctc tattttttttt tttttttttt tttttttttt tttttttttt	6900
taaattttgtt caaatgtttt ttctgtgtctt atcaatatgtt ccatgtgttttt tttttttttt	6960
acctgtgtt gggacaaattt acgtttaattt gttttttttt tttttttttt tttttttttt	7020
cttggaaataaa ttctacttgg ttgtgggttta tttttttttt tttttttttt tttttttttt	7080
gctaatatttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	7140

tctttcttg	taatgtcatt	ttctagttcc	ggtattaagg	taatgctggc	ctagttgaat	7200
gatttaggaa	gtattccctc	tgcttctgtc	ttctgaggtt	ccgcggccgc	ccgtcgaaaa	7260
acaacgtcgt	gactggaaaa	accctggcgt	tacccaaactt	aatcgcccttg	cagcacatcc	7320
cccttgcgc	agctggcgta	atagcgaaga	ggcccgacacc	gatcgccctt	cccaacagtt	7380
gcgcagcctg	aatggcgaat	ggcgctgtat	gcggtatttt	ctccttacgc	atctgtgcgg	7440
tatttcacac	cgcatacgtc	aaagcaacca	tagtacgcgc	cctgtacgg	cgcattaagc	7500
gcggcgggtg	tggtggttac	gcgcagcgtg	accgctacac	ttgccagcgc	cctagcgc	7560
gctccttcg	ctttcttccc	ttccttctc	gccacgttcg	ccggcttcc	ccgtcaagct	7620
ctaaatcggg	ggctccctt	agggttccga	tttagtgctt	tacggcacct	cgaccccaaa	7680
aaacttgatt	tgggtgatgg	ttcacgtagt	ggccatcgc	cctgatagac	ggtttttcgc	7740
ccttgcgt	tggagtccac	gttcttaat	agtggactct	tgttccaaac	tgaaacaaca	7800
ctcaacccta	tctcgggcta	ttctttgtat	ttataaggga	tttgccgat	ttcggcctat	7860
tggtaaaaaa	atgagctgat	ttaacaaaaa	ttaacgoga	attttaacaa	aatattaacg	7920
tttacaattt	tatggtgac	tctcagtaca	atctgctctg	atgcccata	gttaagccag	7980
ccccgacacc	cgccaaacacc	cgctgacgac	ccctgacggg	cttgcgtct	ccggcatcc	8040
gtttacagac	aagctgtgac	cgtctccggg	agctgcgtgt	gtcagaggtt	ttcaccgtca	8100
tcaccgaaac	gcgcgagaga	cgaaaggccc	tcgtgatacg	cctattttta	taggttaatg	8160
tcatgataat	aatggtttct	tagacgtcag	gtggcacttt	tcggggaaat	gtgcgcggaa	8220
cccttatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	8280
cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	8340
tcgcccattat	tccctttttt	gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	8400
tggtaaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	8460
atctcaacag	cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	8520
gcacttttaa	agttctgcta	tgtggcgccg	tattatcccg	tattgacgcc	ggcaagagc	8580
aactcggcgt	ccgcatacac	tatttcaga	atgacttggt	tgagtactca	ccagtcacag	8640
aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgcgtcc	ataaccatga	8700
gtgataacac	tgcggccaac	ttacttctga	caacgatcgg	aggacccaag	gagctaaccg	8760
cttttttgc	caacatgggg	gatcatgtaa	ctcgcccttga	tcgttggaa	ccggagctga	8820

atgaagccat accaaacgac gagcgtgaca ccacgatgcc ttagcaatg gcaacaacgt	8880
tgcgcaaact attaactggc gaactactta ctctagcttc cggcaacaa ttaatagact	8940
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggccttccg gctggctgg	9000
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatcatt gcagcactgg	9060
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta	9120
tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac	9180
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta	9240
aaaggatcta ggtgaagatc cttttgata atctcatgac caaaatccct taacgtgagt	9300
tttcgttcca ctgagcgtca gaccccttag aaaagatcaa aggatcttct tgagatcctt	9360
ttttctgcg cgtaatctgc tgcttgaaa caaaaaacc accgttacca gcggtggtt	9420
gtttgccgga tcaagagcta ccaactctt ttccgaaggt aactggcttc agcagagcgc	9480
agataccaaa tactgttott ctatgttagc cgtatgttcc accacttc aagaactctg	9540
tagcaccgcc tacataccotc gctctgctaa tcctgttacc agtggctgct gccagtggcg	9600
ataagtctgt tcttaccggg ttggactcaa gacgatagtt accggataag ggcgcgg	9660
cgggctgaac ggggggttcg tgcacacagc ccagcttggc gcaacgacc tacaccgaac	9720
tgagataacct acagcgttag ctatgagaaa gcccacgtc tcccaagg agaaaggcgg	9780
acaggtatcc ggtaaaggc agggctggaa caggagagcg cacgagggag ctccagggg	9840
gaaacgcctg gtatctttat agtcctgtcg gtttcgcca cctctgactt gagcgtcgat	9900
ttttgtatg ctcgtcaggg gggcggagcc tatggaaaaa cggccacac gcccctttt	9960
tacggttctt ggcctttgc tggccttttgc ctcacatgtt ctttcgtcg ttatcccctg	10020
attctgtgga taaccgtatt accgcctttg agttagctga taccgtcg cgcagccgaa	10080
cgaccgagcg cagcgagtca gtgagcgagg aagcggaa gcccggggc aaggtcgcca	10140
cgcacaagat caatattaac aatcgtcat ctctcttag caataaaaag gtaaaaatt	10200
acattttaaa aatgacacca tagacgtatgt atgaaaataa tctacttgaa aataaatcta	10260
ggcaaagaag tgcaagactg ttacccagaa aacttacaaa ttgtaaatga gaggttagtg	10320
aagattttaaa tgaatgaaga tctaaataaa cttataaattt gtgagagaaa ttaatgaatg	10380
tctaagttaa tgcagaaacg gagagacata ctatattcat gaactaaaag acttaatatt	10440
gtgaaggtat actttctttt cacataaattt ttagtcaat atgttccaccc caaaaaagct	10500
gtttgttaac ttgtcaacct catttcaaaaa tttatataa aagccaaag acaataacaa	10560

aaatattctt gtagaacaaa atggaaaga atgttccact aaatatcaag atttagagca 10620
aagcatgaga tgtgtggga tagacagtga ggctgataaa atagagtaga gctcagaaac 10680
agacccattg atatatgtaa gtgacctatg aaaaaatata ggcattttac aatgggaaaa 10740
tgcgtatctt ttctttttt agaaaaacag ggaaatataat ttatatgtaa aaaataaaaag 10800
ggAACCCATA TGTCTACCA TACACACAAA AAAATTCCAG TGAATTATAA GTCTAAATGG 10860
agaaggcaaa actttaaatc tttagaaaa taatatagaa gcatgccatc atgacttcag 10920
tgttagagaaa aatttcttat gactcaaagt cctaaccaca aagaaaagat tgttaattag 10980
attgcata gataatgact tattttaaa attaaaaac cattaagaaa agtcaggcca 11040
tagaaatgaca gaaaatattt gcaacacccc agtaaagaga attgtatata gcaattata 11100
aaaagaagtc ttacaaatca gtaaaaaata aaactagaca aaaatttcaa cagatgaaag 11160
agaaactcta aataatcatt acacatgaga aactcaatct cagaaatcag agaactatca 11220
ttgcatacatac actaaatattt agaaatatta aaaggctaag taacatctgt ggcaatattt 11280
atgttatata accttgatata gatgtatga gaacagtact ttacccatg ggtttccctcc 11340
ccaaaccctt accccagtat aaatcatgac aaatatactt taaaaaccat taccctata 11400
ctaacccatgtt ctcctcaaaa ctgtcaaggt catcaaaaat aagaaaagtc tgaggaactg 11460
tcaaaaactaa gaggaaccca aggagacatg agaattataat gtaatgtggc attctgaatg 11520
agatcccaga acagaaaaag aacagtagct aaaaaactaa tgaaatataa ataaagttt 11580
aacttttagtt tttttttaaa aagagtagca ttaacacggc aaagtcattt tcataatttt 11640
cttgaacatt aagtacaagt ctataattaa aaattttta aatgtatgtct ggaacattgc 11700
cagaaacaga agtacagcag ctatctgtgc tgtcgcctaa ctatccatag ctgattggc 11760
taaaaatgaga tacatcaacg ctccctccatg tttttgttt tctttttaaa tgaaaaactt 11820
tatttttaa gaggagtttcc aggttcatag caaaatttggag aggaaggtac attcaagctg 11880
aggaagttttt cctctattcc tagttactg agagattgca tcatgaatgg gtgttaattt 11940
ttgtcaatg ctttttctgt gtctatcaat atgaccatgt gatttcttc tttaacctgt 12000
tgatggaca aattacgtta attgattttc aaacgttcaa ccacccttac atatctggaa 12060
taaattctac ttgggtgtgg tgtatatttt ttgatacatt ctggattct ttttgctaat 12120
attttgttga aatgtttgtt atctttgttc atgagagata ttggctgtt gttttttttt 12180
cttgcataatgtt cattttcttag ttcccggtatt aaggtatgc tggcctagtt gaatgattta 12240

ggaagtattc cctctgcttc tgtcttctga agcggaaagag cgcccaatac gcaaaccgcc 12300
tctcccccgcg cggtggccga ttcattaatg cagctggcac gacaggtttc ccgactggaa 12360
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccccaggc 12420
tttacacttt atgcttccgg ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca 12480
cacaggaaac agctatgaca tgattacgaa ttaa 12514

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.